Multilevel Picard approximations for high-dimensional semilinear second-order PDEs with Lipschitz nonlinearities

Martin Hutzenthaler[1], Arnulf Jentzen[2], Thomas Kruse[3], and Tuan Anh Nguyen[4]

[1] Faculty of Mathematics, University of Duisburg-Essen, Essen, Germany; e-mail: martin.hutzenthaler@uni-due.de

[2] Faculty of Mathematics and Computer Science, University of Münster, Münster, Germany; e-mail: ajentzen@uni-muenster.de

[3] Institute of Mathematics, University of Gießen, Gießen, Germany; e-mail: thomas.kruse@math.uni-giessen.de

[4] Faculty of Mathematics, University of Duisburg-Essen, Essen, Germany; e-mail: tuan.nguyen@uni-due.de

Monday 12th October, 2020

Abstract

The recently introduced full-history recursive multilevel Picard (MLP) approximation methods have turned out to be quite successful in the numerical approximation of solutions of high-dimensional nonlinear PDEs. In particular, there are mathematical convergence results in the literature which prove that MLP approximation methods do overcome the curse of dimensionality in the numerical approximation of nonlinear second-order PDEs in the sense that the number of computational operations of the proposed MLP approximation method grows at most polynomially in both the reciprocal $1/\varepsilon$ of the prescribed approximation accuracy $\varepsilon > 0$ and the PDE dimension $d \in \mathbb{N} = \{1, 2, 3, \ldots\}$. However, in each of the convergence results for MLP approximation methods in the literature it is assumed that the coefficient functions in front of the second-order differential operator are affine linear. In particular, until today there is no result in the scientific literature which proves that any semilinear second-order PDE with a general time horizon and a non affine linear coefficient function in front of the second-order differential operator can be approximated without the curse of dimensionality. It is the key contribution of this article to overcome this obstacle and to propose and analyze a new type of MLP approximation method for semilinear second-order PDEs with possibly nonlinear coefficient functions in front of the second-order differential operators. In particular, the main result of this article proves that this new MLP approximation method does indeed overcome the curse of dimensionality in the numerical approximation of semilinear second-order PDEs.

Contents

1 Introduction

It is a very challenging task in applied mathematics to design and analyze approximation algorithms for high-dimensional nonlinear partial differential equations (PDEs) and this topic of research has been very intensively studied in the scientific literature in the last two decades. Especially, there are two types of approximation methods which have turned out to be quite successful in the numerical approximation of solutions of high-dimensional nonlinear second-order PDEs, namely, (I) deep learning based approximation methods for PDEs (cf., e.g., [1–3, 9–11, 13, 14, 16, 17, 20, 22, 23, 25, 28–31, 33, 41, 43, 46–52, 55, 56]) and (II) full-history recursive multilevel Picard approximation methods for PDEs (cf. [6, 8, 18, 19, 24, 36, 38–40]; in the following we abbreviate *full-history recursive multilevel Picard* as *MLP*). Deep learning based approximation methods for PDEs are, roughly speaking, based on the idea to (Ia) approximate the PDE problem under consideration through a stochastic optimization problem involving deep neural networks as approximations for the solution or the derivatives of the solution of the PDE under consideration and to (Ib) apply stochastic gradient descent methods to approximately solve the resulting stochastic optimization problem. Even though there are a number of encouraging simulation results for deep learning based approximation methods for PDEs in the scientific literature, there are only partial mathematical error analyses in the scientific literature which only partly explain why deep learning based approximation methods for PDEs can approximately solve high-dimensional PDEs (cf., e.g., [12, 21, 26, 27, 30, 37, 42, 45, 53, 55]). In particular, there are no results in the scientific literature which prove that deep learning based approximation methods for PDEs overcome the curse of dimensionality in the sense that the number of computational operations of any deep learning based approximation method grows at most polynomially in both the reciprocal of the prescribed approximation accuracy and the PDE dimension. MLP approximation methods are, roughly speaking, based on the idea to (IIa) reformulate the PDE under consideration as a stochastic fixed point problem with the PDE solution being the fixed point of the stochastic fixed point equation, to (IIb) approximate the fixed point through Banach fixed point iterates (which are also referred to as Picard iterates in the context of integral fixed point equations), and to (IIc) approximate the resulting Banach fixed point iterates through suitable full-history recursive multilevel Monte Carlo approximations. In the case of MLP approximation methods there are both encouraging numerical simulation results (see [8, 19]) and rigorous mathematical results which prove that

MLP approximation methods do indeed overcome the curse of dimensionality in the numerical approximation of nonlinear second-order PDEs (see [5, 6, 18, 24, 36, 38–40]). However, in each of the convergence results for MLP approximation methods in the scientific literature it is assumed that the coefficient functions in front of the second-order differential operator are affine linear. In particular, until today there is no result in the scientific literature which proves that any semilinear second-order PDE with a general time horizon and a non affine linear coefficient function in front of the second-order differential operator can be approximated without the curse of dimensionality.

It is precisely the subject of this article to overcome this obstacle and to propose and analyze a new type of MLP approximation method for semilinear second-order PDEs with possibly nonlinear coefficient functions in front of the second-order differential operators. In particular, the main result of this article, Theorem 4.2 in Section 4 below, proves that this new MLP approximation method overcomes the curse of dimensionality in the numerical approximation of semilinear second-order PDEs in the sense that the number of computational operations of the proposed MLP approximation method grows at most polynomially in both the reciprocal $1/\varepsilon$ of the prescribed approximation accuracy $\varepsilon \in (0,\infty)$ and the PDE dimension $d \in \mathbb{N} = \{1,2,3,\dots\}$. To briefly outline the contribution of this work within this introductory section, we now present in the following result, Theorem 1.1 below, a special case of Proposition 4.1.

Theorem 1.1. *Let $c, T \in [0,\infty)$, $f \in C(\mathbb{R},\mathbb{R})$, for every $d \in \mathbb{N}$ let $u_d \in C^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$, $\mu_d = (\mu_{d,i})_{i\in\{1,2,\dots,d\}} \in C(\mathbb{R}^d,\mathbb{R}^d)$, $\sigma_d = (\sigma_{d,i,j})_{i,j\in\{1,2,\dots,d\}} \in C(\mathbb{R}^d,\mathbb{R}^{d\times d})$ satisfy for all $t\in[0,T]$, $x=(x_1,x_2,\dots,x_d)$, $y=(y_1,y_2,\dots,y_d)\in\mathbb{R}^d$ that*

$$|f(x_1)-f(y_1)| \le c|x_1-y_1|, \quad |u_d(t,x)|^2 + \max_{i,j\in\{1,2,\dots,d\}} (|\mu_{d,i}(0)|+|\sigma_{d,i,j}(0)|) \le c\Big[d^c + \sum_{i=1}^d |x_i|^2\Big], \tag{1}$$

$$|u_d(T,x)-u_d(T,y)|^2 + \sum_{i=1}^d |\mu_{d,i}(x)-\mu_{d,i}(y)|^2 + \sum_{i,j=1}^d |\sigma_{d,i,j}(x)-\sigma_{d,i,j}(y)|^2 \le c\Big[\sum_{i=1}^d |x_i-y_i|^2\Big], \tag{2}$$

$$\text{and} \quad (\tfrac{\partial}{\partial t}u_d)(t,x) + (\tfrac{\partial}{\partial x}u_d)(t,x)\,\mu_d(x) + \tfrac{1}{2}\mathrm{tr}\big(\sigma_d(x)[\sigma_d(x)]^*(\mathrm{Hess}_x\, u)(t,x)\big) = -f(u_d(t,x)), \tag{3}$$

let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\Theta = \bigcup_{n\in\mathbb{N}}\mathbb{Z}^n$, let $\mathfrak{r}^\theta\colon\Omega\to[0,1]$, $\theta\in\Theta$, be i.i.d. random variables[1]*, let $W^{d,\theta}\colon[0,T]\times\Omega\to\mathbb{R}^d$, $d\in\mathbb{N}$, $\theta\in\Theta$, be i.i.d. standard Brownian motions, assume for all $t\in(0,1)$ that $\mathbb{P}(\mathfrak{r}^0\le t)=t$, assume that $(\mathfrak{r}^\theta)_{\theta\in\Theta}$ and $(W^{d,\theta})_{(d,\theta)\in\mathbb{N}\times\Theta}$ are independent, for every $d,N\in\mathbb{N}$, $\theta\in\Theta$, $x\in\mathbb{R}^d$, $t\in[0,T)$ let $Y^{d,N,\theta,x}_t = (Y^{d,N,\theta,x}_{t,s})_{s\in[t,T]}\colon[t,T]\times\Omega\to\mathbb{R}^d$ satisfy for all $n\in\{0,1,\dots,N\}$, $s\in[\frac{nT}{N},\frac{(n+1)T}{N}]\cap[t,T]$ that $Y^{d,N,\theta,x}_{t,t}=x$ and*

$$\begin{split} &Y^{d,N,\theta,x}_{t,s} - Y^{d,N,\theta,x}_{t,\max\{t,nT/N\}} \\ &= \mu_d\big(Y^{d,N,\theta,x}_{t,\max\{t,nT/N\}}\big)\big(s-\max\{t,\tfrac{nT}{N}\}\big) + \sigma_d\big(Y^{d,N,\theta,x}_{t,\max\{t,nT/N\}}\big)\big(W^{d,\theta}_s - W^{d,\theta}_{\max\{t,nT/N\}}\big), \end{split} \tag{4}$$

let $U^{d,\theta}_{n,M}\colon[0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $d,n,M\in\mathbb{Z}$, $\theta\in\Theta$, satisfy for all $d,M\in\mathbb{N}$, $n\in\mathbb{N}_0$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$\begin{split} U^{d,\theta}_{n,M}(t,x) &= \tfrac{\mathbb{1}_{\mathbb{N}}(n)}{M^n}\sum_{i=1}^{M^n} u_d\big(T,Y^{d,M^M,(\theta,0,-i),x}_{t,T}\big) \\ &+ \sum_{\ell=0}^{n-1}\Bigg[\tfrac{(T-t)}{M^{n-\ell}}\sum_{i=1}^{M^{n-\ell}}\big(f\circ U^{d,(\theta,\ell,i)}_{\ell,M} - \mathbb{1}_{\mathbb{N}}(\ell)\,f\circ U^{d,(\theta,-\ell,i)}_{\ell-1,M}\big)\big(t+(T-t)\mathfrak{r}^{(\theta,\ell,i)}, Y^{d,M^M,(\theta,\ell,i),x}_{t,t+(T-t)\mathfrak{r}^{(\theta,\ell,i)}}\big)\Bigg], \end{split} \tag{5}$$

and for every $d,n,M\in\mathbb{N}$ let $\mathfrak{C}_{d,n,M}\in\mathbb{N}$ be the number of function evaluations of f, $u_d(T,\cdot)$, μ_d, and σ_d and the number of realizations of scalar random variables which are used to compute one realization of $U^{d,0}_{n,M}(0,0)\colon\Omega\to\mathbb{R}^d$ (cf. (167) for a precise definition). Then there exist $\mathfrak{c}\in\mathbb{R}$ and $\mathsf{n}\colon\mathbb{N}\times(0,1]\to\mathbb{N}$ such that for all $d\in\mathbb{N}$, $\varepsilon\in(0,1]$ it holds that $\big(\mathbb{E}\big[|u_d(0,0)-U^{d,0}_{\mathsf{n}(d,\varepsilon),\mathsf{n}(d,\varepsilon)}(0,0)|^2\big]\big)^{1/2}\le\varepsilon$ and $\mathfrak{C}_{d,\mathsf{n}(d,\varepsilon),\mathsf{n}(d,\varepsilon)}\le\mathfrak{c}d^{\mathfrak{c}}\varepsilon^{-5}$.

[1] Note that the expression *i.i.d.* is an abbreviation for the expression *independent and identically distributed*.

Theorem 1.1 follows from Corollary 4.4. Corollary 4.4, in turn, follows from Theorem 4.2 (see Section 4 for details). In the following we add a few comments concerning the mathematical objects appearing in Theorem 1.1 above. The real number $T \in (0,\infty)$ in Theorem 1.1 above specifies the time horizon for the PDEs (see (3)) whose solutions we intend to approximate in Theorem 1.1 above. The real number $c \in (0,\infty)$ in Theorem 1.1 above is a constant which we employ to formulate several regularity hypotheses in Theorem 1.1 above. The function $f\colon \mathbb{R} \to \mathbb{R}$ in Theorem 1.1 above describes the nonlinearity for the PDEs (see (3)) whose solutions we intend to approximate in Theorem 1.1 above. The functions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, in Theorem 1.1 above describe the PDE solutions which we intend to approximate in Theorem 1.1 above. The functions $\mu_d\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$, in Theorem 1.1 above describe the coefficient functions in front of the first-order derivative terms in the PDEs (see (3)) whose solutions we intend to approximate in Theorem 1.1 above. The functions $\sigma_d\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $d \in \mathbb{N}$, in Theorem 1.1 above describe the coefficient functions in front of the second-order derivative terms in the PDEs (see (3)) whose solutions we intend to approximate in Theorem 1.1 above. In (1) and (2) we formulate the Lipschitz hypotheses which we employ in Theorem 1.1 above. In (3) we specify the PDEs whose solutions we intend to approximate in Theorem 1.1 above. The probability space $(\Omega,\mathcal{F},\mathbb{P})$ in Theorem 1.1 above is the probability space on which we introduce the stochastic MLP approximations which we employ to approximate the solutions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (3). The set $\Theta = \bigcup_{n\in\mathbb{N}} \mathbb{Z}^n$ in Theorem 1.1 above is used as an index set to introduce sufficiently many independent random variables on this index set. The functions $\mathfrak{r}^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, describe on $[0,1]$ continuously uniformly distributed independent random variables which we use as random input sources for the MLP approximations which we employ in Theorem 1.1 above to approximately compute the solutions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (3). The functions $W^{d,\theta}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $d \in \mathbb{N}$, $\theta \in \Theta$, describe independent standard Brownian motions which we use as random input sources for the MLP approximations which we employ in Theorem 1.1 above to approximately compute the solutions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (3). The functions $Y_t^{d,N,\theta,x}\colon [t,T]\times\Omega \to \mathbb{R}^d$, $d,N \in \mathbb{N}$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, $t \in [0,T)$, in (4) above describe Euler-Mayurama approximations which we use in the MLP approximations in (5) in Theorem 1.1 above as discretizations of the underlying Itô processes associated to the linear parts of the PDEs in (3). The functions $U_{n,M}^{d,\theta}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$, $d,n,M \in \mathbb{N}_0$, $\theta \in \Theta$, in (5) describe the MLP approximations which we employ in Theorem 1.1 above to approximately compute the solutions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (3). The natural numbers $\mathfrak{C}_{d,n,M} \in \mathbb{N}$, $d,n,M \in \mathbb{N}$, describe the sum of the number of function evaluations of f, of the number of function evaluations of $u_d(T,\cdot)$, of the number of function evaluations of μ_d, of the number of function evaluations of σ_d, and of the number of realizations of scalar random variables which are used to compute one realization of the MLP approximations which we employ in Theorem 1.1 above to approximately compute the solutions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (3) (cf. also (167) in Corollary 4.4 in Section 4 for a precise definition of $(\mathfrak{C}_{d,n,M})_{(d,n,M)\in\mathbb{N}^3} \subseteq \mathbb{N}$). Theorem 1.1 establishes that the solutions $u_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (3) can be approximated by the MLP approximations $U_{n,M}^{d,\theta}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$, $d,n,M \in \mathbb{N}_0$, $\theta \in \Theta$, in (5) with the number of involved function evaluations of f, $u_d(T,\cdot)$, μ_d, and σ_d and the number of involved scalar random variables growing at most quintically in the reciprocal $1/\varepsilon$ of the prescribed approximation accuracy $\varepsilon \in (0,\infty)$ and at most polynomially in the PDE dimension $d \in \mathbb{N}$. Our proofs of Theorem 1.1 above and Theorem 4.2 below, respectively, are partially based on previous analyses for MLP approximations in the scientific literature (cf., e.g., [5, 6, 18, 24, 36, 38–40]) and on analyses for numerical approximations for SDEs with non-globally Lipschitz continuous coefficient functions (cf., e.g., [34]).

The remainder of this article is organized as follows. As mentioned above, MLP approximation methods are, roughly speaking, based on the idea to reformulate the PDE under consideration (see (3)) as a stochastic fixed point equation (see (IIa) above) and then to approximate the fixed point of the stochastic fixed point equation through suitable full-history recursive multilevel Monte Carlo approximations (see (IIb) and (IIc) above). In Section 2 below we establish existence, uniqueness, and regularity properties for solutions of such stochastic fixed point equations. In Section 3 below we introduce MLP approximations for solutions of such stochastic fixed point equations (see (44) in Setting 3.1 in Subsection 3.1 below), we study measurability, integrability, and independence properties for the introduced MLP approximations (see Lemma 3.2, Lemma 3.3, and Lemma 3.4 in Subsection 3.2 below), and we establish in Corollary 3.12 in Subsection 3.5 below upper bounds for the L^2-distances between the exact solutions of the considered stochastic fixed point equations and the proposed MLP approximations. In our proof of Corollary 3.12 we employ certain function space-valued Gronwall-type inequalities, which we establish in Lemma 3.9, Lemma 3.10, and Lemma 3.11 in Subsection 3.4 below. In Section 4 we combine the existence, uniqueness, and regularity properties for solutions of stochastic fixed point equations, which we have established in Section 2, with the error analysis for MLP approximations for stochastic fixed point equations, which we have established in Section 3 (see Corollary 3.12 in Subsection 3.5), to obtain a computational complexity analysis for MLP approximations for semilinear second-order PDEs with possibly nonlinear coefficient functions in front of the second-order differential operators.

2 Stochastic fixed-point equations

In this section we establish in the elementary results in Proposition 2.2 and Lemma 2.3 below existence, uniqueness, and regularity properties for solutions of stochastic fixed point equations. Similar existence, uniqueness, and regularity results for solutions of stochastic fixed point equations can, e.g., be found in Beck et al. [4] and Beck et al. [7, Theorem 3.7]. In our proof of Proposition 2.2 we use the well-known auxiliary measurability result in Lemma 2.1 below. For completeness we also include in this section a detailed proof for Lemma 2.1.

2.1 Existence of solutions of stochastic fixed-point equations

Lemma 2.1. *Let $(X,\mathcal{X})$ be a measurable space, let $(Y,\mathcal{Y},\mu)$ be a sigma-finite measure space, let $f\colon X\times Y\to\mathbb{R}$ be measurable, and assume for all $x\in X$ that $\int_Y |f(x,y)|\,\mu(dy)<\infty$. Then it holds that $X\ni x\mapsto\int_Y f(x,y)\,\mu(dy)\in\mathbb{R}$ is measurable.*

Proof of Lemma 2.1. Throughout this proof let $\mathfrak{f}_k\colon X\times Y\to[0,\infty)$, $k\in\{0,1\}$, satisfy for all $k\in\{0,1\}$, $x\in X$, $y\in Y$ that

$$\mathfrak{f}_k(x,y)=\max\{(-1)^k f(x,y),0\}. \tag{6}$$

Note that Fubini's theorem (cf., e.g., Klenke [44, Theorem 14.16]) ensures that for every measurable $g\colon X\times Y\to[0,\infty)$ it holds that $X\ni x\mapsto\int_Y g(x,y)\,\mu(dy)\in[0,\infty]$ is measurable. This proves that for all $k\in\{0,1\}$ it holds that $X\ni x\mapsto\int_Y\mathfrak{f}_k(x,y)\,\mu(dy)\in[0,\infty)$ is measurable. Combining this with the fact that for all $x\in X$ it holds that $\int_Y f(x,y)\,\mu(dy)=\int_Y\mathfrak{f}_0(x,y)\,\mu(dy)-\int_Y\mathfrak{f}_1(x,y)\,\mu(dy)$ implies that $X\ni x\mapsto\int_Y f(x,y)\,\mu(dy)\in\mathbb{R}$ is measurable. The proof of Lemma 2.1 is thus complete. ☐

Proposition 2.2. *Let $d\in\mathbb{N}$, $L,T,c\in[0,\infty)$, $\mathcal{O}\in\mathcal{B}(\mathbb{R}^d)$ satisfy $\mathcal{O}\neq\emptyset$, let $\|\cdot\|\colon\mathbb{R}^d\to[0,\infty)$ be a norm on $\mathbb{R}^d$, let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $X^x_{t,s}\colon\Omega\to\mathcal{O}$, $s\in[t,T]$, $t\in[0,T]$, $x\in\mathcal{O}$, be random variables, assume for every measurable $\psi\colon[0,T]\times\mathcal{O}\to[0,\infty)$ that $\{(t,s)\in[0,T]^2\colon t\leq s\}\times\mathcal{O}\ni(t,s,x)\mapsto\mathbb{E}[\psi(s,X^x_{t,s})]\in[0,\infty]$ is measurable, let $f\colon[0,T]\times\mathcal{O}\times\mathbb{R}\to\mathbb{R}$, $g\colon\mathcal{O}\to\mathbb{R}$, and $V\colon[0,T]\times\mathcal{O}\to(0,\infty)$ be measurable, assume for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathcal{O}$, $v,w\in\mathbb{R}$ that $|f(t,x,0)|\leq cV(t,x)$, $|g(x)|\leq cV(T,x)$, $\mathbb{E}[V(s,X^x_{t,s})]\leq V(t,x)$, and $|f(t,x,v)-f(t,x,w)|\leq L|v-w|$. Then* *(i) there exists a unique measurable $u\colon [0,T]\times\mathcal{O}\to\mathbb{R}$ which satisfies for all $t\in[0,T]$, $x\in\mathcal{O}$ that $\mathbb{E}\big[|g(X^x_{t,T})|\big]+\int_t^T \mathbb{E}\big[|f(s,X^x_{t,s},u(s,X^x_{t,s}))|\big]\,ds+\sup_{y\in\mathcal{O}}\sup_{s\in[0,T]}\frac{|u(s,y)|}{V(s,y)}<\infty$ and*

$$u(t,x)=\mathbb{E}\big[g(X^x_{t,T})\big]+\int_t^T \mathbb{E}\big[f\big(s,X^x_{t,s},u(s,X^x_{t,s})\big)\big]\,ds \tag{7}$$

and

(ii) it holds that for all $t\in[0,T]$ that

$$\sup_{x\in\mathcal{O}}\left(\frac{|u(t,x)|}{V(t,x)}\right)\le\left[\sup_{x\in\mathcal{O}}\left(\frac{|g(x)|}{V(T,x)}\right)+\sup_{x\in\mathcal{O}}\sup_{s\in[t,T]}\left(\frac{|Tf(s,x,0)|}{V(s,x)}\right)\right]e^{L(T-t)}. \tag{8}$$

Proof of Proposition 2.2. Throughout this proof let $\mathcal{V}$ satisfy

$$\mathcal{V}=\left\{u\colon[0,T]\times\mathcal{O}\to\mathbb{R}\colon\left[u\text{ is measurable and }\left[\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{|u(t,x)|}{V(t,x)}\right)<\infty\right]\right]\right\} \tag{9}$$

and let $\|\cdot\|_\lambda\colon\mathcal{V}\to[0,\infty)$, $\lambda\in\mathbb{R}$, satisfy for every $\lambda\in\mathbb{R}$, $v\in\mathcal{V}$ that

$$\|v\|_\lambda=\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{e^{\lambda t}|v(t,x)|}{V(t,x)}\right). \tag{10}$$

Note that (9) and (10) ensure that for all $\lambda\in\mathbb{R}$ it holds that $(\mathcal{V},\|\cdot\|_\lambda)$ is a normed $\mathbb{R}$-vector space. Next we show that $(\mathcal{V},\|\cdot\|_0)$ is an $\mathbb{R}$-Banach space. For this let $v=(v_n)_{n\in\mathbb{N}}\colon\mathbb{N}\to\mathcal{V}$ satisfy

$$\limsup_{N\to\infty}\big(\sup_{n,m\in\mathbb{N}\cap[N,\infty)}\|v_n-v_m\|_0\big)=0. \tag{11}$$

Observe that (11) demonstrates that for all $t\in[0,T]$, $x\in\mathcal{O}$ it holds that $v_n(t,x)\in\mathbb{R}$, $n\in\mathbb{N}$, is a Cauchy sequence. The fact that $(\mathbb{R},|\cdot|)$ is an $\mathbb{R}$-Banach space hence assures that there exists $\phi\colon[0,T]\times\mathcal{O}\to\mathbb{R}$ which satisfies for all $t\in[0,T]$, $x\in\mathcal{O}$ that $\limsup_{n\to\infty}|\phi(t,x)-v_n(t,x)|=0$. Combining this with the fact that for all $n\in\mathbb{N}$ it holds that v_n is measurable proves that ϕ is measurable. Next observe that the fact that for all $t\in[0,T]$, $x\in\mathcal{O}$ it holds that $\limsup_{n\to\infty}|\phi(t,x)-v_n(t,x)|=0$ yields that for all $N\in\mathbb{N}$ it holds that

$$\begin{aligned}
&\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{|\phi(t,x)|}{V(t,x)}\right)=\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{|\lim_{n\to\infty}v_n(t,x)|}{V(t,x)}\right)\\
&\le\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{[\sup_{n\in\mathbb{N}}|v_n(t,x)|]}{V(t,x)}\right)=\sup_{n\in\mathbb{N}}\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{|v_n(t,x)|}{V(t,x)}\right)=\sup_{n\in\mathbb{N}}\|v_n\|_0\\
&\le\left[\sup_{n\in\mathbb{N}\cap[N,\infty)}\|v_n-v_N\|_0\right]+\left[\max_{n\in\{1,2,\dots,N\}}\|v_n\|_0\right]\\
&\le\left[\sup_{n,m\in\mathbb{N}\cap[N,\infty)}\|v_n-v_m\|_0\right]+\left[\max_{n\in\{1,2,\dots,N\}}\|v_n\|_0\right].
\end{aligned} \tag{12}$$

This and (11) imply that

$$\sup_{t\in[0,T]}\sup_{x\in\mathcal{O}}\left(\frac{|\phi(t,x)|}{V(t,x)}\right)\le\sup_{n\in\mathbb{N}}\|v_n\|_0<\infty. \tag{13}$$

Combining this with the fact that ϕ is measurable proves that $\phi \in \mathcal{V}$. In addition, observe that (11) assures that

$$\begin{aligned}\limsup_{n\to\infty} \|\phi - v_n\|_0 &= \limsup_{n\to\infty} \left[\sup_{t\in[0,T]} \sup_{x\in\mathcal{O}} \left(\frac{\lim_{m\to\infty} |v_m(t,x) - v_n(t,x)|}{V(t,x)} \right) \right] \\ &\le \limsup_{n\to\infty} \left[\sup_{t\in[0,T]} \sup_{x\in\mathcal{O}} \sup_{m\in\mathbb{N}\cap[n,\infty)} \left(\frac{|v_m(t,x) - v_n(t,x)|}{V(t,x)} \right) \right] \\ &= \limsup_{n\to\infty} \left[\sup_{m\in\mathbb{N}\cap[n,\infty)} \|v_m - v_n\|_0 \right] = 0.\end{aligned} \tag{14}$$

This demonstrates that $(\mathcal{V}, \|\cdot\|_0)$ is an $\mathbb{R}$-Banach space. Combining this with the fact that for all $\Lambda \in \mathbb{R}$, $\lambda \in [\Lambda, \infty)$, $v \in \mathcal{V}$ it holds that $\|v\|_\Lambda \le \|v\|_\lambda \le e^{(\lambda-\Lambda)T} \|v\|_\Lambda$ shows that for all $\lambda \in \mathbb{R}$ it holds that $(\mathcal{V}, \|\cdot\|_\lambda)$ is an $\mathbb{R}$-Banach space. Moreover, observe that the fact that for all $x \in \mathcal{O}$ it holds that $|g(x)| \le cV(T,x)$ ensures that for all $t \in [0,T]$, $x \in \mathcal{O}$ it holds that

$$\mathbb{E}\big[|g(X^x_{t,T})|\big] \le \left[\sup_{y\in\mathcal{O}} \frac{|g(y)|}{V(T,y)}\right] \mathbb{E}[V(T, X^x_{t,T})] \le \left[\sup_{y\in\mathcal{O}} \left(\frac{|g(y)|}{V(T,y)}\right)\right] V(t,x) < \infty. \tag{15}$$

This implies that

$$[0,T] \times \mathcal{O} \ni (t,x) \mapsto \mathbb{E}\big[g(X^x_{t,T})\big] \in \mathbb{R} \tag{16}$$

is measurable. Furthermore, observe that the triangle inequality and the fact that for all $t \in [0,T]$, $x \in \mathcal{O}$, $v, w \in \mathbb{R}$ it holds that $|f(t,x,v) - f(t,x,w)| \le L|v-w|$ yield that for all $v \in \mathcal{V}$, $t \in [0,T]$, $s \in [t,T]$, $x \in \mathcal{O}$ it holds that

$$\begin{aligned}&\mathbb{E}\big[|f(s, X^x_{t,s}, v(s, X^x_{t,s}))|\big] \le \mathbb{E}\big[|f(s, X^x_{t,s}, 0)| + L|v(s, X^x_{t,s})|\big] \\ &\le \left[\sup_{r\in[s,T]} \sup_{y\in\mathcal{O}} \left(\frac{|f(r,y,0)| + L|v(r,y)|}{V(r,y)} \right) \right] \mathbb{E}\big[V(s, X^x_{t,s})\big] \\ &\le \left[\sup_{r\in[s,T]} \sup_{y\in\mathcal{O}} \left(\frac{|f(r,y,0)| + L|v(r,y)|}{V(r,y)} \right) \right] V(t,x) < \infty.\end{aligned} \tag{17}$$

This implies that for all $v \in \mathcal{V}$ it holds that

$$[0,T] \times [0,T] \times \mathcal{O} \ni (s,t,x) \mapsto \mathbb{1}_{[t,T]}(s)\, \mathbb{E}\big[f\big(\max\{s,t\}, X^x_{t,\max\{s,t\}}, v(, X^x_{t,\max\{s,t\}})\big)\big]\, ds \in \mathbb{R} \tag{18}$$

is measurable. Moreover, observe that (17) implies that for all $v \in \mathcal{V}$, $t \in [0,T]$, $x \in \mathcal{O}$ it holds that

$$\int_0^T \mathbb{1}_{[t,T]}(s) \left|\mathbb{E}\big[f\big(\max\{s,t\}, X^x_{t,\max\{s,t\}}, v(\max\{s,t\}, X^x_{t,\max\{s,t\}})\big)\big]\right| ds < \infty. \tag{19}$$

Lemma 2.1 and (16) hence prove that for all $v \in \mathcal{V}$ it holds that

$$\begin{aligned}&[0,T] \times \mathcal{O} \ni (t,x) \mapsto \mathbb{E}\big[g(X^x_{t,T})\big] \\ &\quad + \int_0^T \mathbb{1}_{[t,T]}(s)\, \mathbb{E}\big[f\big(\max\{s,t\}, X^x_{t,\max\{s,t\}}, v(\max\{s,t\}, X^x_{t,\max\{s,t\}})\big)\big]\, ds \in \mathbb{R}\end{aligned} \tag{20}$$

is measurable. This and (17) imply that there exists $\Phi \colon \mathcal{V} \to \mathcal{V}$ which satisfies for all $t \in [0,T]$, $x \in \mathcal{O}$, $v \in \mathcal{V}$ that

$$(\Phi(v))(t,x) = \mathbb{E}\big[g(X^x_{t,T})\big] + \int_t^T \mathbb{E}\big[f\big(s, X^x_{t,s}, v(s, X^x_{t,s})\big)\big]\, ds. \tag{21}$$

In addition, note that the fact that for all $t \in [0,T]$, $x \in \mathcal{O}$, $v, w \in \mathbb{R}$ it holds that $|f(t,x,v) - f(t,x,w)| \leq L|v-w|$ ensures that for all $\lambda \in (0,\infty)$, $v, w \in \mathcal{V}$ it holds that $\|\Phi(v) - \Phi(w)\|_\lambda \leq \frac{L}{\lambda}\|v-w\|_\lambda$ (cf., e.g., Beck et al. [4, Lemma 2.8]). Hence, we obtain for all $\lambda \in [2L,\infty)$, $v, w \in \mathcal{V}$ that

$$\|\Phi(v) - \Phi(w)\|_\lambda \leq \frac{1}{2}\|v-w\|_\lambda. \tag{22}$$

Banach's fixed point theorem therefore demonstrates that there exists a unique $u \in \mathcal{V}$ which satisfies $\Phi(u) = u$. Combining this, (9), and (17) with (21) establishes item (i). Next observe that (15) and (17) imply that for all $t \in [0,T]$ it holds that

$$\begin{aligned} &\sup_{r\in[t,T]}\sup_{x\in\mathcal{O}}\left(\frac{|u(r,x)|}{V(r,x)}\right) = \sup_{r\in[t,T]}\sup_{x\in\mathcal{O}}\left(\frac{|(\Phi(u))(r,x)|}{V(r,x)}\right) \\ &\leq \sup_{y\in\mathcal{O}}\left(\frac{|g(y)|}{V(T,y)}\right) + \sup_{r\in[0,T]}\sup_{y\in\mathcal{O}}\left(\frac{|Tf(r,y,0)|}{V(r,y)}\right) + L\int_t^T\left[\sup_{r\in[s,T]}\sup_{y\in\mathcal{O}}\left(\frac{|u(r,y)|}{V(r,y)}\right)\right]ds. \end{aligned} \tag{23}$$

This, the fact that $u \in \mathcal{V}$, and Gronwall's lemma yield that for all $t \in [0,T]$ it holds that

$$\sup_{r\in[t,T]}\sup_{x\in\mathcal{O}}\left(\frac{|u(r,x)|}{V(r,x)}\right) \leq \left[\sup_{y\in\mathcal{O}}\left(\frac{|g(y)|}{V(T,y)}\right) + \sup_{r\in[0,T]}\sup_{y\in\mathcal{O}}\left(\frac{|Tf(r,y,0)|}{V(r,y)}\right)\right]e^{L(T-t)}. \tag{24}$$

This establishes item (ii). The proof of Proposition 2.2 is thus complete. □

2.2 Perturbation analysis for stochastic fixed-point equations

Lemma 2.3. *Let $d \in \mathbb{N}$, $g \in C(\mathbb{R}^d,\mathbb{R})$, $c, L, \rho, \eta \in [0,\infty)$, $T, \delta, p, q \in (0,\infty)$ satisfy $p^{-1} + q^{-1} \leq 1$, let $\|\cdot\|\colon \mathbb{R}^d \to [0,\infty)$ be a seminorm on $\mathbb{R}^d$, let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $X^{x,k}_{t,(\cdot)} = (X^{x,k}_{t,s}(\omega))_{(s,\omega)\in[t,T]\times\Omega}\colon [t,T]\times\Omega \to \mathbb{R}^d$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $k \in \{1,2\}$, be measurable, let $f\colon [0,T]\times\mathbb{R}^d\times\mathbb{R} \to \mathbb{R}$, $V\colon \mathbb{R}^d \to (0,\infty)$, $\psi\colon [0,T]\times\mathbb{R}^d \to (0,\infty)$, and $u_k\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $k \in \{1,2\}$, be measurable, assume for all $s \in [0,T]$, $r \in [s,T]$ and all measurable $h\colon \mathbb{R}^d\times\mathbb{R}^d \to [0,\infty)$ that $\mathbb{R}^d\times\mathbb{R}^d \ni (y_1,y_2) \mapsto \mathbb{E}\big[h\big(X^{y_1,1}_{s,r}, X^{y_2,1}_{s,r}\big)\big] \in [0,\infty]$ is measurable, and assume for all $x, y \in \mathbb{R}^d$, $t \in [0,T]$, $s \in [t,T]$, $r \in [s,T]$, $v, w \in \mathbb{R}$, $k \in \{1,2\}$ and all measurable $h\colon \mathbb{R}^d\times\mathbb{R}^d \to [0,\infty)$ that*

$$X^{x,k}_{t,t} = x, \qquad \mathbb{E}\big[\psi(s,X^{x,k}_{t,s})\big] \leq \eta\psi(t,x), \tag{25}$$

$$\mathbb{E}\Big[\mathbb{E}\big[h\big(X^{\mathfrak{x},1}_{s,r}, X^{\mathfrak{y},1}_{s,r}\big)\big]\big|_{(\mathfrak{x},\mathfrak{y})=(X^{x,1}_{t,s},X^{y,1}_{t,s})}\Big] = \mathbb{E}\big[h\big(X^{x,1}_{t,r}, X^{y,1}_{t,r}\big)\big], \tag{26}$$

$$\max\{T|f(t,x,v) - f(t,y,w)|, |g(x) - g(y)|\} \leq LT|v-w| + T^{-1/2}[V(x)+V(y)]^{1/p}\|x-y\|, \tag{27}$$

$$\mathbb{E}\Big[\mathbb{E}\big[\|X^{\mathfrak{x},1}_{s,r} - X^{\mathfrak{y},1}_{s,r}\|^q\big]\big|_{(\mathfrak{x},\mathfrak{y})=(X^{x,1}_{t,s},X^{x,2}_{t,s})}\Big] \leq \delta^q\psi(t,x), \tag{28}$$

$$\mathbb{E}\big[|g(X^{x,k}_{t,T})|\big] + \int_t^T \mathbb{E}\big[|f(s,X^{x,k}_{t,s},u_k(s,X^{x,k}_{t,s}))|\big]\,ds < \infty, \tag{29}$$

$$u_k(t,x) = \mathbb{E}\left[g(X^{x,k}_{t,T}) + \int_t^T f(s,X^{x,k}_{t,s},u_k(s,X^{x,k}_{t,s}))\,ds\right], \tag{30}$$

and $\max\{|u_k(t,x)|^p, ce^{\rho(t-s)}\mathbb{E}[V(X^{x,k}_{t,s})]\} \leq cV(x)$. Then it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$|u_1(t,x) - u_2(t,x)| \leq 4(1+LT)T^{-1/2}\exp\!\big((L + \tfrac{\rho}{p} + \eta^{1/q}L)(T-t)\big)|V(x)|^{1/p}|\psi(t,x)|^{1/q}\delta. \tag{31}$$

Proof of Lemma 2.3. Throughout this proof assume w.l.o.g. that $c > 0$. Note that Hölder's inequality implies that for all $s \in [0,T]$, $r \in [s,T]$, $x_1, x_2 \in \mathbb{R}^d$, $k_1, k_2 \in \{1,2\}$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\big(V(X^{x_1,k_1}_{s,r}) + V(X^{x_2,k_2}_{s,r})\big)^{1/p} \big\|X^{x_1,k_1}_{s,r} - X^{x_2,k_2}_{s,r}\big\|\Big] \\
&\le \big(\mathbb{E}\big[V(X^{x_1,k_1}_{s,r}) + V(X^{x_2,k_2}_{s,r})\big]\big)^{1/p} \big(\mathbb{E}\big[\big\|X^{x_1,k_1}_{s,r} - X^{x_2,k_2}_{s,r}\big\|^q\big]\big)^{1/q} \\
&\le e^{\rho(r-s)/p}(V(x_1) + V(x_2))^{1/p} \big(\mathbb{E}\big[\big\|X^{x_1,k_1}_{s,r} - X^{x_2,k_2}_{s,r}\big\|^q\big]\big)^{1/q}.
\end{aligned} \tag{32}$$

This, (30), the triangle inequality, Tonelli's theorem, and (25)–(27) show that for all $t \in [0,T]$, $s \in [t,T]$, $x_1, x_2 \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[|u_1(s, X^{x_1,1}_{t,s}) - u_1(s, X^{x_2,1}_{t,s})|\big] \\
&= \mathbb{E}\Bigg[\bigg|\mathbb{E}\Big[g(X^{y_1,1}_{s,T}) - g(X^{y_2,1}_{s,T}) \\
&\quad + \int_s^T f(r, X^{y_1,1}_{s,r}, u_1(r, X^{y_1,1}_{s,r})) - f(r, X^{y_2,1}_{s,r}, u_1(r, X^{y_2,1}_{s,r}))\, dr\Big]\bigg|\bigg|_{(y_1,y_2)=(X^{x_1,1}_{t,s}, X^{x_2,1}_{t,s})}\Bigg] \\
&\le \mathbb{E}\big[|g(X^{x_1,1}_{t,T}) - g(X^{x_2,1}_{t,T})|\big] \\
&\quad + \int_s^T \mathbb{E}\big[\big|f(r, X^{x_1,1}_{t,r}, u_1(r, X^{x_1,1}_{t,r})) - f(r, X^{x_2,1}_{t,r}, u_1(r, X^{x_2,1}_{t,r}))\big|\big]\, dr \\
&\le \mathbb{E}\Big[T^{-1/2}\big(V(X^{x_1,1}_{t,T}) + V(X^{x_2,1}_{t,T})\big)^{1/p} \big\|X^{x_1,1}_{t,T} - X^{x_2,1}_{t,T}\big\|\Big] \\
&\quad + L\int_s^T \mathbb{E}\big[\big|u_1(r, X^{x_1,1}_{t,r}) - u_1(r, X^{x_2,1}_{t,r})\big|\big]\, dr \\
&\quad + \int_s^T \mathbb{E}\Big[T^{-3/2}\big(V(X^{x_1,1}_{t,r}) + V(X^{x_2,1}_{t,r})\big)^{1/p} \big\|X^{x_1,1}_{t,r} - X^{x_2,1}_{t,r}\big\|\Big]\, dr \\
&\le L\int_s^T \mathbb{E}\big[\big|u_1(r, X^{x_1,1}_{t,r}) - u_1(r, X^{x_2,1}_{t,r})\big|\big]\, dr + e^{\rho(T-t)/p}(V(x_1) + V(x_2))^{1/p} \\
&\quad \cdot \Bigg[T^{-1/2}\Big(\mathbb{E}\Big[\big\|X^{x_1,1}_{t,T} - X^{x_2,1}_{t,T}\big\|^q\Big]\Big)^{1/q} + \int_t^T T^{-3/2}\Big(\mathbb{E}\Big[\big\|X^{x_1,1}_{t,r} - X^{x_2,1}_{t,r}\big\|^q\Big]\Big)^{1/q} dr\Bigg].
\end{aligned} \tag{33}$$

This, Gronwall's lemma, and the fact that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that $\max\{e^{-\rho(s-t)}\mathbb{E}\big[V(X^{x,k}_{t,s})\big], |u_1(t,x)|^p\} \le cV(x)$ imply that for all $t \in [0,T]$, $s \in [t,T]$, $x_1, x_2 \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[|u_1(s, X^{x_1,1}_{t,s}) - u_1(s, X^{x_2,1}_{t,s})|\big] \le e^{\rho(T-t)/p}(V(x_1) + V(x_2))^{1/p} \\
&\cdot \Bigg[T^{-1/2}\Big(\mathbb{E}\Big[\big\|X^{x_1,1}_{t,T} - X^{x_2,1}_{t,T}\big\|^q\Big]\Big)^{1/q} + \int_t^T T^{-3/2}\Big(\mathbb{E}\Big[\big\|X^{x_1,1}_{t,r} - X^{x_2,1}_{t,r}\big\|^q\Big]\Big)^{1/q} dr\Bigg] e^{L(T-s)}.
\end{aligned} \tag{34}$$

Combining this with (25) assures that for all $s \in [0,T]$, $x_1, x_2 \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&|u_1(s, x_1) - u_1(s, x_2)| \le e^{(L+\rho/p)(T-s)}(V(x_1) + V(x_2))^{1/p} \\
&\cdot \Bigg[T^{-1/2}\Big(\mathbb{E}\Big[\big\|X^{x_1,1}_{s,T} - X^{x_2,1}_{s,T}\big\|^q\Big]\Big)^{1/q} + \int_s^T T^{-3/2}\big(\mathbb{E}\big[\big\|X^{x_1,1}_{s,r} - X^{x_2,1}_{s,r}\big\|^q\big]\big)^{1/q} dr\Bigg].
\end{aligned} \tag{35}$$

This, Hölder's inequality, (26), and (28) assure that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that

$$
\begin{aligned}
&\mathbb{E}\big[|u_1(s,X^{x,1}_{t,s})-u_1(s,X^{x,2}_{t,s})|\big] \\
&\le e^{(L+\rho/p)(T-s)}\mathbb{E}\Bigg[\Big[T^{-1/2}(V(z_1)+V(z_2))^{1/p}\left(\mathbb{E}\Big[\big\|X^{z_1,1}_{s,T}-X^{z_2,1}_{s,T}\big\|^q\Big]\right)^{1/q}\Big]\Big|_{(z_1,z_2)=(X^{x,1}_{t,s},X^{x,2}_{t,s})}\Bigg] \\
&+e^{(L+\rho/p)(T-s)}\int_s^T\mathbb{E}\Big[\Big[T^{-3/2}(V(z_1)+V(z_2))^{1/p}\left(\mathbb{E}\big[\big\|X^{z_1,1}_{s,r}-X^{z_2,1}_{s,r}\big\|^q\big]\right)^{1/q}\Big]\Big|_{(z_1,z_2)=(X^{x,1}_{t,s},X^{x,2}_{t,s})}\Big]\,dr \\
&\le 2e^{(L+\rho/p)(T-s)}T^{-1/2}\left(\mathbb{E}[V(X^{x,1}_{t,s})+V(X^{x,2}_{t,s})]\right)^{1/p} \\
&\cdot\sup_{r\in[s,T]}\left(\mathbb{E}\Big[\mathbb{E}\Big[\big\|X^{z_1,1}_{s,r}-X^{z_2,1}_{s,r}\big\|^q\Big]\Big|_{(z_1,z_2)=(X^{x,1}_{t,s},X^{x,2}_{t,s})}\Big]\right)^{1/q} \\
&\le 2T^{-1/2}e^{(L+\rho/p)(T-t)}(2V(x))^{1/p}\delta(\psi(t,x))^{1/q}.
\end{aligned}
\tag{36}
$$

This, (30), the triangle inequality, (27), (32), and (28) show that for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $\varepsilon\in(0,1)$ it holds that

$$
\begin{aligned}
&|u_1(t,x)-u_2(t,x)| \\
&=\left|\mathbb{E}\Big[g(X^{x,1}_{t,T})-g(X^{x,2}_{t,T})+\int_t^T f(s,X^{x,1}_{t,s},u_1(s,X^{x,1}_{t,s}))-f(s,X^{x,2}_{t,s},u_2(s,X^{x,2}_{t,s}))\,ds\Big]\right| \\
&\le\mathbb{E}\big[|g(X^{x,1}_{t,T})-g(X^{x,2}_{t,T})|\big]+\int_t^T\mathbb{E}\big[|f(s,X^{x,1}_{t,s},u_1(s,X^{x,1}_{t,s})-(f(s,X^{x,2}_{t,s},u_1(s,X^{x,2}_{t,s}))|\big]\,ds \\
&\quad+\int_t^T\mathbb{E}\big[|f(s,X^{x,2}_{t,s},u_1(s,X^{x,2}_{t,s}))-f(s,X^{x,2}_{t,s},u_2(s,X^{x,2}_{t,s}))|\big]\,ds \\
&\le\mathbb{E}\Big[T^{-1/2}\left(V(X^{x,1}_{t,T})+V(X^{x,2}_{t,T})\right)^{1/p}\big\|X^{x,1}_{t,T}-X^{x,2}_{t,T}\big\|\Big] \\
&\quad+\int_t^T\mathbb{E}\Big[T^{-3/2}\left(V(X^{x,1}_{t,s})+V(X^{x,2}_{t,s})\right)^{1/p}\big\|X^{x,1}_{t,s}-X^{x,2}_{t,s}\big\|\Big]\,ds \\
&\quad+L\int_t^T\mathbb{E}\big[|u_1(s,X^{x,1}_{t,s})-u_1(s,X^{x,2}_{t,s})|\big]\,ds+L\int_t^T\mathbb{E}\big[|u_1(s,X^{x,2}_{t,s})-u_2(s,X^{x,2}_{t,s})|\big]\,ds \\
&\le T^{-1/2}e^{\rho(T-t)/p}(2V(x))^{1/p}\delta(\varepsilon+\psi(t,x))^{1/q}+T^{-1/2}e^{\rho(T-t)/p}(2V(x))^{1/p}\delta(\varepsilon+\psi(t,x))^{1/q} \\
&\quad+LT\left[2T^{-1/2}e^{(L+\rho/p)(T-t)}(2V(x))^{1/p}\delta(\varepsilon+\psi(t,x))^{1/q}\right] \\
&\quad+L\int_t^T\left[\sup_{r\in[s,T]}\sup_{z\in\mathbb{R}^d}\frac{e^{\rho r/p}|u_1(r,z)-u_2(r,z)|}{(V(z))^{1/p}(\varepsilon+\psi(r,z))^{1/q}}\right]e^{-\rho s/p}\mathbb{E}\big[(V(X^{x,2}_{t,s}))^{1/p}(\varepsilon+\psi(s,X^{x,2}_{t,s}))^{1/q}\big]\,ds.
\end{aligned}
\tag{37}
$$

Next note that (25) implies for all $t\in[0,T]$, $x\in\mathbb{R}^d$ that $\psi(t,x)\le\eta\psi(t,x)$. The fact that $\psi>0$ hence demonstrates that $\eta\ge1$. Hölder's inequality, the assumption that $1/p+1/q\le1$, and (25) therefore assure that for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$, $\varepsilon\in(0,1)$ it holds that

$$
\begin{aligned}
&\mathbb{E}\big[(V(X^{x,2}_{t,s})^{1/p}(\varepsilon+\psi(X^{x,2}_{t,s}))^{1/q}\big]\le\left(\mathbb{E}[V(X^{x,2}_{t,s})]\right)^{1/p}\left(\varepsilon+\mathbb{E}[\psi(s,X^{x,2}_{t,s})]\right)^{1/q} \\
&\le e^{\rho(s-t)/p}(V(x))^{1/p}(\varepsilon+\eta\psi(t,x))^{1/q}\le e^{\rho(s-t)/p}(V(x))^{1/p}\left[\eta(\varepsilon+\psi(t,x))\right]^{1/q}.
\end{aligned}
\tag{38}
$$

This and (37) demonstrate that for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $\varepsilon\in(0,1)$ it holds that

$$
\begin{aligned}
&|u_1(t,x)-u_2(t,x)|\le(2+2LT)\left[T^{-1/2}e^{(L+\rho/p)(T-t)}(2V(x))^{1/p}\delta(\varepsilon+\psi(t,x))^{1/q}\right] \\
&+L\int_t^T\left[\sup_{r\in[s,T],z\in\mathbb{R}^d}\frac{e^{\rho r/p}|u_1(r,z)-u_2(r,z)|}{(V(z))^{1/p}(\varepsilon+\psi(r,z))^{1/q}}\right]e^{-\rho s/p}e^{\rho(s-t)/p}(V(x))^{1/p}\eta^{1/q}(\varepsilon+\psi(t,x))^{1/q}\,ds.
\end{aligned}
\tag{39}
$$

This and the fact $2^{1/p} \leq 2$ imply that for all $t \in [0,T]$, $\varepsilon \in (0,1)$ it holds that

$$\begin{aligned}&\left[\sup_{r\in[t,T],x\in\mathbb{R}^d} \frac{e^{(L+\rho/p)r}|u_1(r,x)-u_2(r,x)|}{(V(x))^{1/p}(\varepsilon+\psi(r,x))^{1/q}}\right]\\&\leq 4(1+LT)T^{-1/2}e^{(L+\rho/p)T}\delta + L\eta^{1/q}\int_t^T\left[\sup_{r\in[s,T],z\in\mathbb{R}^d}\frac{e^{(L+\rho/p)r}|u_1(r,z)-u_2(r,z)|}{(V(z))^{1/p}(\varepsilon+\psi(r,z))^{1/q}}\right]ds.\end{aligned} \tag{40}$$

Gronwall's lemma therefore ensures that for all $t \in [0,T]$, $\varepsilon \in (0,1)$ it holds that

$$\left[\sup_{r\in[t,T],x\in\mathbb{R}^d} \frac{e^{(L+\rho/p)r}|u_1(r,x)-u_2(r,x)|}{(V(x))^{1/p}(\varepsilon+\psi(r,x))^{1/q}}\right] \leq 4(1+LT)T^{-1/2}e^{(L+\rho/p)T}\delta e^{L\eta^{1/q}(T-t)}. \tag{41}$$

Hence, we obtain that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\varepsilon \in (0,1)$ it holds that

$$|u_1(t,x)-u_2(t,x)| \leq 4(1+LT)T^{-1/2}e^{(L+\rho/p+L\eta^{1/q})(T-t)}(V(x))^{1/p}(\varepsilon+\psi(t,x))^{1/q}\delta. \tag{42}$$

This completes the proof of Lemma 2.3. ☐

3 Full-history recursive multilevel Picard (MLP) approximations

In this section we introduce MLP approximations for solutions of stochastic fixed point equations (see (44) in Setting 3.1 in Subsection 3.1 below), we study measurability, integrability, and independence properties for the introduced MLP approximations (see Lemma 3.2, Lemma 3.3, and Lemma 3.4 in Subsection 3.2 below), and we establish in Corollary 3.12 in Subsection 3.5 below upper bounds for the L^2-distances between the exact solutions of the considered stochastic fixed point equations and the proposed MLP approximations. In our proof of Corollary 3.12 we employ certain function space-valued Gronwall-type inequalities, which we establish in Lemma 3.9, Lemma 3.10, and Lemma 3.11 in Subsection 3.4 below. Our proof of Lemma 3.9 employs the well-known and elementary auxiliary results in Lemma 3.6, Lemma 3.7, and Lemma 3.8. For completeness we include in this section also detailed proofs for Lemma 3.6, Lemma 3.7, and Lemma 3.8.

3.1 Mathematical description of MLP approximations

Setting 3.1. *Let $d \in \mathbb{N}$, $T,c,L,\rho \in [0,\infty)$, $\Delta = \{(t,s) \in [0,T]^2 : t \leq s\}$, let $f\colon [0,T]\times\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$, $g\colon\mathbb{R}^d\to\mathbb{R}$, $u\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$, and $\varphi\colon\mathbb{R}^d\to(0,\infty)$ be measurable, let $F\colon\mathbb{R}^{[0,T]\times\mathbb{R}^d}\to\mathbb{R}^{[0,T]\times\mathbb{R}^d}$ satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $y,\mathfrak{y}\in\mathbb{R}$, $w\in\mathbb{R}^{[0,T]\times\mathbb{R}^d}$ that*

$$|f(t,x,y)-f(t,x,\mathfrak{y})| \leq L|y-\mathfrak{y}| \qquad \text{and} \qquad (F(w))(t,x) = f(t,x,w(t,x)), \tag{43}$$

let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\Theta = \bigcup_{n\in\mathbb{N}}\mathbb{Z}^n$, let $\mathfrak{r}^\theta\colon\Omega\to[0,1]$, $\theta\in\Theta$, be i.i.d. random variables, let $Y^\theta = (Y^\theta_{t,s}(x,\omega))_{(t,s,x,\omega)\in\Delta\times\mathbb{R}^d\times\Omega}\colon\Delta\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$, $\theta\in\Theta$, be measurable, assume for all $t\in(0,1)$ that $\mathbb{P}(\mathfrak{r}^0\leq t)=t$, let $\mathcal{R}^\theta\colon[0,T]\times\Omega\to[0,T]$, $\theta\in\Theta$, satisfy for all $t\in[0,T]$, $\theta\in\Theta$ that $\mathcal{R}^\theta_t = t+(T-t)\mathfrak{r}^\theta$, assume for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ that $Y^\theta_{t,s}(x)$, $\theta\in\Theta$, are i.i.d., assume that $(\mathfrak{r}^\theta)_{\theta\in\Theta}$ and $(Y^\theta_{t,s}(x))_{(\theta,t,s,x)\in\Theta\times\Delta\times\mathbb{R}^d}$ are independent, let $U^\theta_{n,M}\colon[0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $n,M\in\mathbb{Z}$, $\theta\in\Theta$, satisfy for all $n\in\mathbb{N}_0$, $M\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$\begin{aligned}U^\theta_{n,M}(t,x) &= \frac{\mathbb{1}_{\mathbb{N}}(n)}{M^n}\sum_{i=1}^{M^n} g\big(Y^{(\theta,0,-i)}_{t,T}(x)\big)\\&+\sum_{\ell=0}^{n-1}\frac{(T-t)}{M^{n-\ell}}\left[\sum_{i=1}^{M^{n-\ell}}\big(F(U^{(\theta,\ell,i)}_{\ell,M})-\mathbb{1}_{\mathbb{N}}(\ell)F(U^{(\theta,-\ell,i)}_{\ell-1,M})\big)\big(\mathcal{R}^{(\theta,\ell,i)}_t,Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\big)\right],\end{aligned} \tag{44}$$

assume for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ *that* $\mathbb{E}\big[|g(Y^0_{t,T}(x))|\big] + \int_t^T \mathbb{E}\big[|(F(u))(r, Y^0_{t,r}(x))|\big]\, dr < \infty$, $\mathbb{E}\big[\varphi(Y^0_{t,s}(x))\big] \le e^{\rho(s-t)}\varphi(x)$, $|(F(0))(t,x)|^2 + |g(x)|^2 + |u(t,x)|^2 \le c\varphi(x)$, *and*

$$u(t,x) = \mathbb{E}\left[g(Y^0_{t,T}(x)) + \int_t^T \big(F(u)\big)\big(r, Y^0_{t,r}(x)\big)\, dr \right]. \tag{45}$$

3.2 Measurability, integrability, and independence properties for MLP approximations

Lemma 3.2 (Independence and distributional properties)**.** *Assume Setting 3.1. Then*

(i) it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ *that* $U^\theta_{n,M}$ *and* $F(U^\theta_{n,M})$ *are measurable,*

(ii) it holds[2] *for all* $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ *that*

$$\mathfrak{S}\big((U^\theta_{n,M}(t,x))_{(t,x)\in[0,T]\times\mathbb{R}^d}\big) \subseteq \mathfrak{S}\big((\mathfrak{r}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y^{(\theta,\vartheta)}_{t,s}(x))_{(\vartheta,t,s,x)\in\Theta\times\Delta\times\mathbb{R}^d}\big), \tag{46}$$

(iii) it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ *that* $(U^\theta_{n,M}(t,x))_{(t,x)\in[0,T]\times\mathbb{R}^d}$, $(Y^\theta_{t,s}(x))_{(t,s,x)\in\Delta\times\mathbb{R}^d}$, *and* $\mathfrak{r}^\theta$ *are independent,*

(iv) it holds for all $n, m \in \mathbb{N}_0$, $M \in \mathbb{N}$, $i, j, k, \ell, \nu \in \mathbb{Z}$, $\theta \in \Theta$ *with* $(i,j) \ne (k,l)$ *that* $(U^{(\theta,i,j)}_{n,M}(t,x))_{(t,x)\in[0,T]\times\mathbb{R}^d}$, $(U^{(\theta,k,\ell)}_{m,M}(t,x))_{(t,x)\in[0,T]\times\mathbb{R}^d}$, $\mathfrak{r}^{(\theta,i,j)}$, *and* $(Y^{(\theta,i,j)}_{t,s}(x))_{(t,s,x)\in\Delta\times\mathbb{R}^d}$ *are independent, and*

(v) it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ *that* $U^\theta_{n,M}(t,x)$, $\theta \in \Theta$, *are identically distributed.*

Proof of Lemma 3.2. Note that (44), (43), the assumption that for all $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U^\theta_{0,M} = 0$, the assumption that f is measurable, the assumption that g is measurable, the assumption that for all $\theta \in \Theta$ it holds that Y^θ is measurable, the fact that for all $\theta \in \Theta$ it holds that $\mathcal{R}^\theta$ is measurable, and induction establish item (i). Next observe that (43), item (i), and the assumption that f is measurable assure that for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $F(U^\theta_{n,M})$ is measurable. The assumption that for all $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U^\theta_{0,M} = 0$, (44), the assumption that for all $\theta \in \Theta$ it holds that Y^θ is measurable, the fact that for all $\theta \in \Theta$ it holds that $\mathcal{R}^\theta$ is measurable, and induction hence prove item (ii). Furthermore, note that item (ii) and the fact that for all $\theta \in \Theta$ it holds that $(\mathfrak{r}^{(\theta,\vartheta)})_{\vartheta\in\Theta}$, $(Y^{(\theta,\vartheta)}_{t,s}(x))_{(\vartheta,t,s,x)\in\Theta\Delta\times\mathbb{R}^d}$, $(Y^\theta_{t,s}(x))_{(t,s,x)\in\Delta\times\mathbb{R}^d}$, and $\mathfrak{r}^\theta$ are independent establish item (iii). In addition, note that item (ii), the fact that for all $i, j \in \mathbb{Z}$, $\theta \in \Theta$ it holds that $\mathfrak{r}^{(\theta,i,j)}$ and $(Y^{(\theta,i,j)}_{t,s}(x))_{(t,s,x)\in\Delta\times\mathbb{R}^d}$ are independent, and the fact that for all $i, j, k, \ell \in \mathbb{Z}$, $\theta \in \Theta$ with $(i,j) \ne (k,\ell)$ it holds that

$$\big(\mathfrak{r}^{(\theta,i,j,\vartheta)}, Y^{(\theta,i,j,\vartheta)}_{t,s}(x)\big)_{(\vartheta,t,s,x)\in\Theta\times\Delta\times\mathbb{R}^d} \tag{47}$$

and

$$\big(\mathfrak{r}^{(\theta,k,\ell,\vartheta)}, Y^{(\theta,k,\ell,\vartheta)}_{t,s}(x)\big)_{(\vartheta,t,s,x)\in\Theta\times\Delta\times\mathbb{R}^d} \tag{48}$$

are independent prove item (iv). Furthermore, note that the assumption that for all $M \in \mathbb{N}, \theta \in \Theta$ it holds that $U^\theta_{0,M} = 0$, the assumption that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that $Y^\theta_{t,s}(x)$, $\theta \in \Theta$, are i.i.d., the fact that for all $t \in [0,T]$ it holds that $\mathcal{R}^\theta_t$, $\theta \in \Theta$, are i.i.d., items (i)–(iv), induction, and, e.g., [38, Lemma 2.4] establish item (v). The proof of Lemma 3.2 is thus complete. □

[2] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $n \in \mathbb{N}$, and let $(S_k, \mathcal{S}_k)$, $k \in \{1, 2, ..., n\}$, be measurable spaces. Note that for all $X_k\colon \Omega \to S_k$, $k \in \{1, 2, ..., n\}$, it holds that $\mathfrak{S}(X_1, X_2, ..., X_n)$ is the smallest sigma-algebra on Ω with respect to which $X_1, X_2, ..., X_n$ are measurable.

Lemma 3.3 (Integrability)**.** *Assume Setting 3.1, let $M \in \mathbb{N}$, and let $\dim\colon \Theta \to \mathbb{N}$ satisfy for all $n \in \mathbb{N}$, $\theta \in \mathbb{Z}^n$ that $\dim(\theta) = n$. Then*

(i) it holds for all $t \in [0,T]$, $\ell \in \mathbb{N}_0$, $\eta, \mu, \nu \in \Theta$ with $\min\{\dim(\eta), \dim(\mu)\} \geq \dim(\nu)$ that

$$\sup_{x \in \mathbb{R}^d} \left[(\varphi(x))^{-1} e^{\rho t} \mathbb{E}\Big[\Big| (T-t)\, \big(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{\mu}_{\ell-1,M})\big)(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \Big|^2 \Big] \right]^{1/2}$$
$$\leq \left[\sup_{s \in [0,T]} \sup_{x \in \mathbb{R}^d} \Big(\mathbb{1}_{\{0\}}(\ell)(T-t)(\varphi(x))^{-1/2} e^{\rho s/2} |(F(0))(s,x)| \Big) \right]$$
$$+ \left[\int_t^T \sup_{r \in [s,T]} \sup_{x \in \mathbb{R}^d} \left[\mathbb{1}_{\mathbb{N}}(\ell)(T-t)^{1/2} L (\varphi(x))^{-1} e^{\rho r} \mathbb{E}\Big[\big| U^{\eta}_{\ell,M}(r,x) - U^{\nu}_{\ell-1,M}(r,x) \big|^2 \Big] \right] ds \right]^{1/2}, \tag{49}$$

(ii) it holds for all $\theta \in \Theta$ that

$$\sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \big[(\varphi(x))^{-1} e^{\rho t} \mathbb{E}\big[|g(Y^{\theta}_{t,T}(x))|^2 \big] \big] \leq \sup_{x \in \mathbb{R}^d} \big[(\varphi(x))^{-1} |g(x)|^2 \big] e^{\rho T} < \infty, \tag{50}$$

(iii) it holds for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $\eta, \mu, \nu \in \Theta$ with $\min\{\dim(\eta), \dim(\mu)\} \geq \dim(\nu)$ that

$$\sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left[\frac{e^{\rho t}}{\varphi(x)} \mathbb{E}\Big[\big| (T-t)(F(U^{\eta}_{n,M}) - \mathbb{1}_{\mathbb{N}}(n) F(U^{\mu}_{n-1,M}))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \big|^2 \Big] \right]^{1/2} < \infty, \tag{51}$$

(iv) it holds for all $n \in \mathbb{N}_0$, $\theta \in \Theta$ that

$$\sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left[(\varphi(x))^{-1} e^{\rho t} \mathbb{E}\Big[\big| U^{\theta}_{n,M}(t,x) \big|^2 \Big] \right]^{1/2} < \infty, \tag{52}$$

and

(v) it holds for all $n \in \mathbb{N}_0$, $\eta, \nu \in \Theta$ with $\dim(\eta) \geq \dim(\nu)$ that

$$\sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left[\frac{e^{\rho t}}{\varphi(x)} \mathbb{E}\Big[\big| (T-t)(F(U^{\eta}_{n,M}))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \big|^2 \Big] \right]^{1/2}$$
$$= \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left[\frac{e^{\rho t}(T-t)}{\varphi(x)} \int_t^T \mathbb{E}\Big[\big| (F(U^{\eta}_{n,M}))(s, Y^{\nu}_{t,s}(x)) \big|^2 \Big] ds \right]^{1/2} < \infty. \tag{53}$$

Proof of Lemma 3.3. Observe that item (ii) in Lemma 3.2 and the assumption that $(\mathfrak{r}^{\theta})_{\theta \in \Theta}$ and $(Y^{\theta}_{t,s}(x))_{(\theta,t,s,x) \in \Theta \times \Delta \times \mathbb{R}^d}$ are independent show that for all $\ell \in \mathbb{N}_0$, $\eta, \mu, \nu \in \Theta$ with $\min\{\dim(\eta), \dim(\mu)\} \geq \dim(\nu)$ it holds that $((F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{\mu}_{\ell-1,M}))(t,x))_{(t,x) \in [0,T] \times \mathbb{R}^d}$, $\mathfrak{r}^{\nu}$, and $(Y^{\nu}_{t,s}(x))_{(t,s,x) \in \Delta \times \mathbb{R}^d}$ are independent. Combining item (i) in Lemma 3.2, the assumption that for all $\theta \in \Theta$ it holds that Y^{θ} is measurable, the fact that for all $\nu \in \Theta$, $r \in (0,1)$ it holds that $\mathbb{P}(\mathfrak{r}^{\nu} \leq r) = r$, and the fact that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that $\mathbb{E}\big[\varphi(Y^0_{t,s}(x))\big] \leq e^{\rho(s-t)} \varphi(x)$ with, e.g., [38, Lemma 2.2] therefore implies that for all $\ell \in \mathbb{N}_0$, $t \in [0,T)$, $x \in \mathbb{R}^d$, $\eta, \mu, \nu \in \Theta$ with $\min\{\dim(\eta), \dim(\mu)\} \geq \dim(\nu)$ it holds that

$$\mathbb{E}\Big[\Big| (T-t)\, \big(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{\mu}_{\ell-1,M})\big)\, (\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \Big|^2 \Big]$$
$$= \frac{1}{T-t} \int_t^T \mathbb{E}\Big[\mathbb{E}\Big[\big| (T-t)\, \big(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{\mu}_{\ell-1,M})\big)\, (s,z) \big|^2 \Big] \Big|_{z = Y^{\nu}_{t,s}(x)} \Big] ds \tag{54}$$
$$= (T-t) \int_t^T \mathbb{E}\Big[\mathbb{E}\Big[\big| \big(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{\mu}_{\ell-1,M})\big)\, (s,z) \big|^2 \Big] \Big|_{z = Y^{\nu}_{t,s}(x)} \Big] ds,$$

$$\mathbb{E}\Big[\big|(T-t)(F(U^{\eta}_{n,M}))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x))\big|^2\Big] = (T-t)\int_t^T \mathbb{E}\Big[\big|(F(U^{\eta}_{n,M}))(s, Y^{\nu}_{t,s}(x))\big|^2\Big]\,ds, \qquad (55)$$

and

$$\begin{aligned}
&\mathbb{E}\bigg[\Big|(T-t)\left(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell)F(U^{\mu}_{\ell-1,M})\right)(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x))\Big|^2\bigg] \\
&\leq (T-t)\int_t^T \mathbb{E}\left[\left(\sup_{r\in[s,T],z\in\mathbb{R}^d}\left[\frac{e^{\rho r}}{\varphi(z)}\mathbb{E}\Big[\big|F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell)F(U^{\mu}_{\ell-1,M})(r,z)\big|^2\Big]\right]\right)\frac{\varphi(Y^{\nu}_{t,s}(x))}{e^{\rho s}}\right] ds \\
&\leq (T-t)\int_t^T \left(\sup_{r\in[s,T],z\in\mathbb{R}^d}\left[\frac{e^{\rho r}}{\varphi(z)}\mathbb{E}\Big[\big|F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell)F(U^{\mu}_{\ell-1,M})(r,z)\big|^2\Big]\right]\right)\frac{e^{\rho(s-t)}\varphi(x)}{e^{\rho s}}\, ds \\
&= e^{-\rho t}\varphi(x)(T-t)^2\left[\frac{1}{T-t}\int_t^T \sup_{r\in[s,T],z\in\mathbb{R}^d}\left[\frac{e^{\rho r}}{\varphi(z)}\mathbb{E}\Big[\big|F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell)F(U^{\mu}_{\ell-1,M})(r,z)\big|^2\Big]\right] ds\right].
\end{aligned} \qquad (56)$$

This, the fact that $\forall\, \eta \in \Theta, t \in [0,T], x \in \mathbb{R}^d\colon U^{\eta}_{0,M}(t,x) = 0$, and (43) imply that for all $\ell \in \mathbb{N}_0$, $t \in [0,T)$, $\eta,\mu,\nu \in \Theta$ with $\min\{\dim(\eta),\dim(\mu)\} \geq \dim(\nu)$ it holds that

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d}\left[(\varphi(x))^{-1}e^{\rho t}\mathbb{E}\bigg[\Big|(T-t)\left(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell)F(U^{\mu}_{\ell-1,M})\right)(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x))\Big|^2\bigg]\right]^{1/2} \\
&\leq (T-t)\left[\sup_{s\in[0,T]}\sup_{x\in\mathbb{R}^d}\left(\mathbb{1}_{\{0\}}(\ell)(\varphi(x))^{-1/2}e^{\rho s/2}|(F(0))(s,x)|\right)\right] \\
&+ (T-t)\left[\frac{1}{T-t}\int_t^T \sup_{r\in[s,T]}\sup_{x\in\mathbb{R}^d}\left[\mathbb{1}_{\mathbb{N}}(\ell)L(\varphi(x))^{-1}e^{\rho r}\mathbb{E}\Big[\big|(U^{\eta}_{\ell,M} - U^{\nu}_{\ell-1,M})(r,x)\big|^2\Big]\right] ds\right]^{1/2} \qquad (57) \\
&= (T-t)\left[\sup_{s\in[0,T]}\sup_{x\in\mathbb{R}^d}\left(\mathbb{1}_{\{0\}}(\ell)(\varphi(x))^{-1/2}e^{\rho s/2}|(F(0))(s,x)|\right)\right] \\
&+ (T-t)^{1/2}\left[\int_t^T \sup_{r\in[s,T]}\sup_{x\in\mathbb{R}^d}\left[\mathbb{1}_{\mathbb{N}}(\ell)L(\varphi(x))^{-1}e^{\rho r}\mathbb{E}\Big[\big|(U^{\eta}_{\ell,M} - U^{\nu}_{\ell-1,M})(r,x)\big|^2\Big]\right] ds\right]^{1/2}.
\end{aligned}$$

This establishes item (i). Next observe that the fact that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that $Y^{\theta}_{t,s}(x)$, $\theta \in \Theta$, are identically distributed and the fact that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that $\mathbb{E}\big[\varphi(Y^0_{t,s}(x))\big] \leq e^{\rho(s-t)}\varphi(x)$ imply that for all $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
\mathbb{E}\Big[\big|g(Y^{\theta}_{t,T}(x))\big|^2\Big] = \mathbb{E}\Big[\big|g(Y^{0}_{t,T}(x))\big|^2\Big] &\leq \mathbb{E}\Big[\sup_{z\in\mathbb{R}^d}\left[|g(z)|^2/\varphi(z)\right]\big|\varphi(Y^0_{t,T}(x))\big|\Big] \\
&\leq \sup_{z\in\mathbb{R}^d}\left[|g(z)|^2/\varphi(z)\right]e^{\rho(T-t)}\varphi(x).
\end{aligned} \qquad (58)$$

This and the fact that for all $y \in \mathbb{R}^d$ it holds that $|g(y)|^2 \leq C\varphi(y)$ imply that for all $\theta \in \Theta$ it holds that

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left[(\varphi(x))^{-1}e^{\rho t}(\mathbb{E}\Big[\big|g(Y^{\theta}_{t,T}(x))\big|^2\Big]\right] \leq \sup_{z\in\mathbb{R}^d}\left[|g(z)|^2/\varphi(z)\right]e^{\rho T} < \infty. \qquad (59)$$

This establishes item (ii). In the next step we prove items (iii) and (iv) by induction on $n \in \mathbb{N}_0$. The fact that $\forall\, \theta \in \Theta\colon U^{\theta}_{0,M} = U^{\theta}_{-1,M} = 0$, item (i), and the fact that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $|(F(0))(t,x)|^2 \leq C\varphi(x)$ show that for all $\nu \in \Theta$ it holds that

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left[(\varphi(x))^{-1}e^{\rho t}\mathbb{E}\Big[\big|(T-t)(F(0))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x))\big|^2\Big]\right]^{1/2} < \infty. \qquad (60)$$

This establishes items (iii) and (iv) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni (n-1) \dashrightarrow n \in \mathbb{N}$ let $n \in \mathbb{N}$ satisfy for all $\ell \in [0, n-1] \cap \mathbb{N}_0$, $\theta \in \Theta$, $\eta, \mu, \nu \in \Theta$ with $\min\{\dim(\eta), \dim(\mu)\} \geq \dim(\nu)$ that

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[(\varphi(x))^{-1} e^{\rho t} \mathbb{E}\left[\left| (T-t)(F(U^{\eta}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{\mu}_{\ell-1,M}))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \right|^2 \right] \right]^{1/2} < \infty \tag{61}$$

and

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[(\varphi(x))^{-1} e^{\rho t} \mathbb{E}\left[\left| U^{\theta}_{\ell,M}(t,x) \right|^2 \right] \right]^{1/2} < \infty. \tag{62}$$

Observe that the triangle inequality, (44), item (ii), and (61) imply that for all $\theta \in \Theta$ it holds that

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[(\varphi(x))^{-1} e^{\rho t} \mathbb{E}\left[\left| U^{\theta}_{n,M}(t,x) \right|^2 \right] \right]^{1/2} < \infty. \tag{63}$$

Combining item (i) and (62) with the fact that $n \in \mathbb{N}$ hence shows that for all $\eta, \mu, \nu \in \Theta$ with $\min\{\dim(\eta), \dim(\mu)\} \geq \dim(\nu)$ it holds that

$$\begin{aligned} &\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[e^{\rho t} (\varphi(x))^{-1} \mathbb{E}\left[\left| (T-t) \left(F(U^{\eta}_{n,M}) - \mathbb{1}_{\mathbb{N}}(n) F(U^{\mu}_{n-1,M}) \right) (\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}}(x)) \right|^2 \right] \right]^{1/2} \\ &\leq TL \left[\sup_{s\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[(\varphi(x))^{-1} e^{\rho s} \mathbb{E}\left[\left| (U^{\eta}_{n,M} - U^{\nu}_{n-1,M})(s,x) \right|^2 \right] \right] \right]^{1/2} < \infty. \end{aligned} \tag{64}$$

Induction, (62), and (63) hence establish items (iii) and (iv). Next observe that the triangle inequality and (61) ensure that for all $n \in \mathbb{N}_0$, $\eta, \nu \in \Theta$ with $\dim(\eta) \geq \dim(\nu)$ it holds that

$$\begin{aligned} &\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[\frac{e^{\rho t}}{\varphi(x)} \mathbb{E}\left[\left| (T-t)(F(U^{\eta}_{n,M}))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \right|^2 \right] \right]^{1/2} \\ &\leq \sum_{l=0}^{n} \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[\tfrac{e^{\rho t}(T-t)^2}{\varphi(x)} \mathbb{E}\left[\left| (F(U^{\eta}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l) F(U^{\eta}_{l-1,M})))(\mathcal{R}^{\nu}_t, Y^{\nu}_{t,\mathcal{R}^{\nu}_t}(x)) \right|^2 \right] \right]^{1/2} < \infty. \end{aligned} \tag{65}$$

This and (55) establish item (v). The proof of Lemma 3.3 is thus complete. ◻

Lemma 3.4 (Expectations of approximations)**.** *Assume Setting 3.1 and let $\theta \in \Theta$. Then*

(i) it holds for all $\ell \in \mathbb{N}_0$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\big(F(U^{(\theta,\ell,i)}_{\ell,M}) - \mathbb{1}_{\mathbb{N}}(\ell) F(U^{(\theta,-\ell,i)}_{\ell-1,M})\big)\big(\mathcal{R}^{(\theta,\ell,i)}_t, Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\big)$, $i \in \mathbb{N}$, are i.i.d. and

(ii) it holds for all $n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} \mathbb{E}\big[U^{\theta}_{n,M}(t,x)\big] &= \mathbb{E}\big[g\left(Y^{\theta}_{t,T}(x)\right)\big] + (T-t)\mathbb{E}\Big[\big(F(U^{\theta}_{n-1,M})\big)\Big(\mathcal{R}^{\theta}_t, Y^{\theta}_{t,\mathcal{R}^{\theta}_t}(x)\Big)\Big] \\ &= \mathbb{E}\big[g\left(Y^{\theta}_{t,T}(x)\right)\big] + \int_t^T \mathbb{E}\big[\big(F(U^{\theta}_{n-1,M})\big)\big(s, Y^{\theta}_{t,s}(x)\big)\big]\, ds. \end{aligned} \tag{66}$$

Proof of Lemma 3.4. Observe that item (ii) in Lemma 3.2 shows that for all $i \in \mathbb{N}$, $\ell \in \mathbb{N}_0$ it holds that

$$\mathfrak{S}\Big((U^{(\theta,\ell,i)}_{\ell,M}(t,x))_{(t,x)\in[0,T]\times\mathbb{R}^d} \Big) \subseteq \mathfrak{S}\Big((\mathfrak{r}^{(\theta,\ell,i,\vartheta)})_{\vartheta\in\Theta}, (Y^{(\theta,\ell,i,\vartheta)}_{s,t}(x))_{(\vartheta,s,t,x)\in\Theta\times\Delta\times\mathbb{R}^d} \Big) \tag{67}$$

and

$$\mathfrak{S}\Big(\big(U^{(\theta,-\ell,i)}_{\ell-1,M}(t,x)\big)_{(t,x)\in[0,T]\times\mathbb{R}^d}\Big)\subseteq\mathfrak{S}\Big(\big(\mathfrak{r}^{(\theta,-\ell,i,\vartheta)}\big)_{\vartheta\in\Theta},\big(Y^{(\theta,-\ell,i,\vartheta)}_{s,t}(x)\big)_{(\vartheta,s,t,x)\in\Theta\times\Delta\times\mathbb{R}^d}\Big).\tag{68}$$

Combining the fact that $\mathfrak{r}^\nu$, $\nu\in\Theta$, are independent and the fact that for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ it holds that $Y^\nu_{t,s}(x)$, $\nu\in\Theta$, are independent hence assures that for all $\ell\in\mathbb{N}_0$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that $\big(F(U^{(\theta,\ell,i)}_{\ell,M})-\mathbb{1}_{\mathbb{N}}(\ell)F(U^{(\theta,-\ell,i)}_{\ell-1,M})\big)\big(\mathcal{R}^{(\theta,\ell,i)}_t,Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\big)$, $i\in\mathbb{N}$, are independent. Moreover, observe that (68), the fact that $\mathfrak{r}^\nu$, $\nu\in\Theta$, are independent, the fact that for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ it holds that $Y^\nu_{t,s}(x)$, $\nu\in\Theta$, are independent, and the fact that $(\mathfrak{r}^\nu)_{\nu\in\Theta}$ and $(Y^\nu_{t,s}(x))_{(\nu,t,s,x)\in\Theta\times\Delta\times\mathbb{R}^d}$ are independent demonstrate that for all $\ell\in\mathbb{N}_0$, $i\in\mathbb{N}$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that $U^{(\theta,\ell,i)}_{\ell,M}$, $U^{(\theta,-\ell,i)}_{\ell-1,M}$, $\mathfrak{r}^{(\theta,\ell,i)}$, and $Y^{(\theta,\ell,i)}$ are independent. Combining this and, e.g., the disintegration-type result in [38, Lemma 2.2] establishes item (i). Next note that item (iii) in Lemma 3.2 implies that for all $\ell\in\mathbb{N}_0$, $k\in\{0,1\}$ it holds that $(U^\theta_{\ell-k,M}(t,x))_{(t,x)\in[0,T]\times\mathbb{R}^d}$, $\mathfrak{r}^\theta$, and $(Y^\theta_{t,s}(x))_{(t,s,x)\in\{(r,u)\in[0,T]^2:u\in[r,T]\}\times\mathbb{R}^d}$ are independent. Combining the fact that for all $\ell\in\mathbb{N}_0$, $i\in\mathbb{N}$, $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ it holds that $U^{(\theta,\ell,i)}(t,x)$ and $U^\theta(t,x)$ are identically distributed (see item (v) in Lemma 3.2), the fact that for all $\ell\in\mathbb{N}_0$, $i\in\mathbb{N}$ it holds that $\mathfrak{r}^{(\theta,\ell,i)}$ and $\mathfrak{r}^\theta$ are identically distributed, the fact that for all $\ell\in\mathbb{N}_0$, $i\in\mathbb{N}$, $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ it holds that $Y^{(\theta,\ell,i)}_{t,s}(x)$ and $Y^\theta_{t,s}(x)$ are identically distributed, item (iii) in Lemma 3.3, and, e.g., the disintegration-type result in [38, Lemma 2.2] hence proves that for all $i\in\mathbb{N}$, $\ell\in\mathbb{N}_0$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\Big(F(U^{(\theta,\ell,i)}_{\ell,M})-\mathbb{1}_{\mathbb{N}}(\ell)F(U^{(\theta,-\ell,i)}_{\ell-1,M})\Big)\Big(\mathcal{R}^{(\theta,\ell,i)}_t,Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\Big)\Big]\\
&=\mathbb{E}\Big[\Big(F(U^{(\theta,\ell,i)}_{\ell,M})\Big)\Big(\mathcal{R}^{(\theta,\ell,i)}_t,Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\Big)\Big]-\mathbb{1}_{\mathbb{N}}(\ell)\mathbb{E}\Big[\Big(F(U^{(\theta,-\ell,i)}_{\ell-1,M})\Big)\Big(\mathcal{R}^{(\theta,\ell,i)}_t,Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\Big)\Big]\\
&=\mathbb{E}\Big[\Big(F(U^\theta_{\ell,M})\Big)\Big(\mathcal{R}^\theta_t,Y^\theta_{t,\mathcal{R}^\theta_t}(x)\Big)\Big]-\mathbb{1}_{\mathbb{N}}(\ell)\mathbb{E}\Big[\Big(F(U^\theta_{\ell-1,M})\Big)\Big(\mathcal{R}^\theta_t,Y^\theta_{t,\mathcal{R}^\theta_t}(x)\Big)\Big].
\end{aligned}\tag{69}$$

The assumption that for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ it holds that $Y^\vartheta_{t,s}(x)$, $\vartheta\in\Theta$, are identically distributed, item (iii) in Lemma 3.3, item (iii) in Lemma 3.2, the fact that for all $t\in[0,T)$ it holds that $\mathcal{R}_t$ is continuous uniformly distributed on $[t,T]$, and, e.g., the disintegration-type result in [38, Lemma 2.2] therefore imply that for all $n\in\mathbb{N}$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}
\mathbb{E}\big[U^\theta_{n,M}(t,x)\big]&=\frac{1}{M^n}\sum_{i=1}^{M^n}\mathbb{E}\Big[g\Big(Y^{(\theta,0,-i)}_{t,T}(x)\Big)\Big]\\
&+\sum_{\ell=0}^{n-1}\frac{(T-t)}{M^{n-\ell}}\left[\sum_{i=1}^{M^{n-\ell}}\mathbb{E}\Big[\Big(F(U^{(\theta,\ell,i)}_{\ell,M})-\mathbb{1}_{\mathbb{N}}(\ell)F(U^{(\theta,-\ell,i)}_{\ell-1,M})\Big)\Big(\mathcal{R}^{(\theta,\ell,i)}_t,Y^{(\theta,\ell,i)}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}(x)\Big)\Big]\right]
\end{aligned}\tag{70}$$

and

$$\begin{aligned}
&\mathbb{E}\big[U^\theta_{n,M}(t,x)\big]=\mathbb{E}\big[g\big(Y^\theta_{t,T}(x)\big)\big]\\
&\quad+(T-t)\sum_{\ell=0}^{n-1}\left[\mathbb{E}\Big[\big(F(U^\theta_{\ell,M})\big)\Big(\mathcal{R}^\theta_t,Y^\theta_{t,\mathcal{R}^\theta_t}(x)\Big)\Big]-\mathbb{1}_{\mathbb{N}}(\ell)\mathbb{E}\Big[\big(F(U^\theta_{\ell-1,M})\big)\Big(\mathcal{R}^\theta_t,Y^\theta_{t,\mathcal{R}^\theta_t}(x)\Big)\Big]\right]\\
&=\mathbb{E}\big[g\,\big(Y^\theta_{t,T}(x)\big)\big]+(T-t)\mathbb{E}\Big[\big(F(U^\theta_{n-1,M})\big)\Big(\mathcal{R}^\theta_t,Y^\theta_{t,\mathcal{R}^\theta_t}(x)\Big)\Big]\\
&=\mathbb{E}\big[g\,\big(Y^\theta_{t,T}(x)\big)\big]+\int_t^T\mathbb{E}\big[\big(F(U^\theta_{n-1,M})\big)\big(s,Y^\theta_{t,s}(x)\big)\big]\,ds.
\end{aligned}\tag{71}$$

This establishes item (ii). The proof of Lemma 3.4 is thus complete. □

3.3 Recursive error bounds for MLP approximations

Lemma 3.5 (Error recursion). *Assume Setting 3.1 and let $n, M \in \mathbb{N}$, $t \in [0,T]$. Then*

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d}\left[e^{\rho t}\varphi(x)^{-1}\mathbb{E}\left[\left|U^0_{n,M}(t,x)-u(t,x)\right|^2\right]\right]^{1/2}\\
&\leq \frac{2e^{\rho T/2}}{\sqrt{M^n}}\sup_{s\in[0,T]}\sup_{z\in\mathbb{R}^d}\left[\frac{\max\{|T(F(0))(s,z)|,|g(z)|\}}{\sqrt{\varphi(z)}}\right]\\
&\quad+\sum_{\ell=0}^{n-1}\left[\frac{2(T-t)^{1/2}L}{\sqrt{M^{n-\ell-1}}}\left(\int_t^T\sup_{r\in[s,T]}\sup_{x\in\mathbb{R}^d}\left[(\varphi(x))^{-1}e^{\rho r}\mathbb{E}\left[\left|U^0_{\ell,M}(r,x)-u(r,x)\right|^2\right]\right]ds\right)^{1/2}\right].
\end{aligned} \tag{72}$$

Proof of Lemma 3.5. Observe that the triangle inequality, (44), Bienaymé's identity, the fact that for all $x \in \mathbb{R}^d$ it holds that $Y^\theta_{t,T}(x)$, $\theta \in \Theta$, are i.i.d., and item (i) in Lemma 3.4 imply that for all $x \in \mathbb{R}^d$ it holds[3] that

$$\begin{aligned}
&\left(\mathbb{V}\mathrm{ar}\big(U^0_{n,M}(t,x)\big)\right)^{1/2} \leq \left(\mathbb{V}\mathrm{ar}\left(\frac{1}{M^n}\sum_{i=1}^{M^n}g\left(Y^{(0,0,-i)}_{t,T}(x)\right)\right)\right)^{1/2}\\
&+\sum_{\ell=0}^{n-1}\left(\mathbb{V}\mathrm{ar}\left(\frac{1}{M^{n-\ell}}\sum_{i=1}^{M^{n-\ell}}(T-t)\left(F\big(U^{(0,\ell,i)}_{\ell,M}\big)-\mathbb{1}_{\mathbb{N}}(\ell)F\big(U^{(0,-\ell,i)}_{\ell-1,M}\big)\right)\Big(\mathcal{R}^{(0,\ell,i)}_t,Y^{(0,\ell,i)}_{t,\mathcal{R}^{(0,\ell,i)}_t}(x)\Big)\right)\right)^{1/2}\\
&\leq \frac{1}{\sqrt{M^n}}\left(\mathbb{E}\left[\left|g(Y^0_{t,T}(x))\right|^2\right]\right)^{1/2}\\
&+\sum_{\ell=0}^{n-1}\frac{1}{\sqrt{M^{n-\ell}}}\left(\mathbb{E}\left[\left|(T-t)\left(F\big(U^{(0,\ell,1)}_{\ell,M}\big)-\mathbb{1}_{\mathbb{N}}(\ell)F\big(U^{(0,-\ell,1)}_{\ell-1,M}\big)\right)\Big(\mathcal{R}^{(0,\ell,1)}_t,Y^{(0,\ell,1)}_{t,\mathcal{R}^{(0,\ell,1)}_t}(x)\Big)\right|^2\right]\right)^{1/2}.
\end{aligned} \tag{73}$$

Lemma 3.3 (applied for every $\ell \in [0, n-1] \cap \mathbb{N}$ with $\eta \curvearrowleft (0,\ell,1)$, $\mu \curvearrowleft (0,-\ell,1)$, $\nu \curvearrowleft (0,\ell,1)$ in the notation of Lemma 3.3) hence shows that for all $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
\left(\mathbb{V}\mathrm{ar}\big(U^0_{n,M}(t,x)\big)\right)^{1/2} &\leq \frac{2e^{\rho(T-t)/2}\sqrt{\varphi(x)}}{\sqrt{M^n}}\sup_{s\in[0,T]}\sup_{z\in\mathbb{R}^d}\left[\frac{\max\{|T(F(0))(s,z)|,|g(z)|\}}{\sqrt{\varphi(z)}}\right]\\
&\quad+\sum_{\ell=1}^{n-1}\frac{e^{-\rho t/2}\sqrt{\varphi(x)}}{\sqrt{M^{n-\ell}}}(T-t)^{1/2}L\\
&\quad\cdot\left(\int_t^T\sup_{r\in[s,T]}\sup_{z\in\mathbb{R}^d}\left[\frac{e^{\rho r}}{\varphi(z)}\mathbb{E}\left[\left|U^{(0,\ell,1)}_{\ell,M}(r,z)-U^{(0,-\ell,1)}_{\ell-1,M}(r,z)\right|^2\right]\right]ds\right)^{1/2}.
\end{aligned} \tag{74}$$

Next note that item (v) in Lemma 3.2 and the triangle inequality demonstrate that for all $\ell \in \mathbb{N}$, $\eta, \nu \in \Theta$, $s \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\left(\mathbb{E}\left[\left|U^\eta_{\ell,M}(s,x)-U^\nu_{\ell-1,M}(s,x)\right|^2\right]\right)^{1/2}\\
&\leq\left(\mathbb{E}\left[\left|U^\eta_{\ell,M}(s,x)-u(s,x)\right|^2\right]\right)^{1/2}+\left(\mathbb{E}\left[\left|U^\nu_{\ell-1,M}(s,x)-u(s,x)\right|^2\right]\right)^{1/2}\\
&=\left(\mathbb{E}\left[\left|U^0_{\ell,M}(s,x)-u(s,x)\right|^2\right]\right)^{1/2}+\left(\mathbb{E}\left[\left|U^0_{\ell-1,M}(s,x)-u(s,x)\right|^2\right]\right)^{1/2}.
\end{aligned} \tag{75}$$

[3] Note that for every probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and every random variable $X\colon \Omega \to \mathbb{R}$ with $\mathbb{E}[|X|] < \infty$ it holds that $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}[|X - \mathbb{E}[X]|^2]$.

This, (74), and the fact that for all $a_0, a_1, \ldots, a_n \in [0, \infty]$ it holds that $\sum_{\ell=1}^{n-1}(a_\ell + a_{\ell-1}) \leq \sum_{\ell=0}^{n-1}[(2 - \mathbb{1}_{\{n-1\}}(\ell))a_\ell]$ imply that for all $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
\Big(\mathbb{V}\mathrm{ar}\big(U^0_{n,M}(t,x)\big)\Big)^{1/2} &\leq \frac{2e^{\rho(T-t)/2}\sqrt{\varphi(x)}}{\sqrt{M^n}} \sup_{s\in[0,T]} \sup_{z\in\mathbb{R}^d} \left[\frac{\max\{|T(F(0))(s,z)|, |g(z)|\}}{\sqrt{\varphi(z)}}\right] \\
&+ \sum_{\ell=0}^{n-1} \Bigg[\frac{(2 - \mathbb{1}_{\{n-1\}}(\ell))e^{-\rho t/2}\sqrt{\varphi(x)}}{\sqrt{M^{n-\ell-1}}}(T-t)^{1/2}L \\
&\cdot \left(\int_t^T \sup_{r\in[s,T]} \sup_{z\in\mathbb{R}^d} \left[\frac{e^{\rho r}\mathbb{E}\big[|U^0_{\ell,M}(r,z) - u(r,z)|^2\big]}{\varphi(z)}\right] ds\right)^{1/2}\Bigg].
\end{aligned} \tag{76}$$

Hence, we obtain that

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d}\Big((\varphi(x))^{-1}e^{\rho t}\mathbb{V}\mathrm{ar}\big(U^0_{n,M}(t,x)\big)\Big)^{1/2} &\leq \frac{2e^{\rho T/2}}{\sqrt{M^n}} \sup_{s\in[0,T]} \sup_{z\in\mathbb{R}^d} \left[\frac{\max\{|T(F(0))(s,z)|, |g(z)|\}}{\sqrt{\varphi(z)}}\right] \\
+ \sum_{\ell=0}^{n-1} \Bigg[\frac{(2 - \mathbb{1}_{\{n-1\}}(\ell))(T-t)^{1/2}L}{\sqrt{M^{n-\ell-1}}} &\left(\int_t^T \sup_{r\in[s,T]} \sup_{x\in\mathbb{R}^d} \left[\frac{e^{\rho r}\mathbb{E}\big[|U^0_{\ell,M}(r,x) - u(r,x)|^2\big]}{\varphi(x)}\right] ds\right)^{1/2}\Bigg].
\end{aligned} \tag{77}$$

Next observe that (45) and item (ii) in Lemma 3.4 imply that for all $x \in \mathbb{R}^d$ it holds that

$$\mathbb{E}\big[U^0_{n,M}(t,x)\big] - u(t,x) = \int_t^T \mathbb{E}\big[(F(U^0_{n-1,M}))(s, Y^0_{t,s}(x)) - (F(u))(s, Y^0_{t,s}(x))\big]\, ds. \tag{78}$$

Combining this, Jensen's inequality, item (iii) in Lemma 3.2, the fact that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that $\mathbb{E}\big[\varphi(Y^0_{t,s}(x))\big] \leq e^{\rho(s-t)}\varphi(x)$, (43), and, e.g., the disintegration-type result in [38, Lemma 2.2] demonstrates that for all $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\big|\mathbb{E}\big[U^0_{n,M}(t,x)\big] - u(t,x)\big| \\
&\leq (T-t)^{1/2}\left(\int_t^T \mathbb{E}\Big[\big|(F(U^0_{n-1,M}))(s, Y^0_{t,s}(x)) - (F(u))(s, Y^0_{t,s}(x))\big|^2\Big]\, ds\right)^{1/2} \\
&= (T-t)^{1/2}\left(\int_t^T \mathbb{E}\Big[\mathbb{E}\Big[\big|(F(U^0_{n-1,M}))(s,z) - (F(u))(s,z)\big|^2\Big]\Big|_{z=Y^0_{t,s}(x)}\Big]\, ds\right)^{1/2} \\
&\leq (T-t)^{1/2}\left(\int_t^T \left[\sup_{r\in[s,T]} \sup_{z\in\mathbb{R}^d} \frac{\mathbb{E}\Big[\big|(F(U^0_{n-1,M}))(r,z) - (F(u))(r,z)\big|^2\Big]}{\varphi(z)}\right] \mathbb{E}\big[\varphi(Y^0_{t,s}(x))\big]\, ds\right)^{1/2} \\
&\leq L(T-t)^{1/2}\left(\int_t^T \left[\sup_{r\in[s,T]} \sup_{z\in\mathbb{R}^d} \frac{\mathbb{E}\Big[\big|U^0_{n-1,M}(r,z) - u(r,z)\big|^2\Big]}{\varphi(z)}\right] \mathbb{E}\big[\varphi(Y^0_{t,s}(x))\big]\, ds\right)^{1/2} \\
&\leq L(T-t)^{1/2}\left(\int_t^T \left[\sup_{r\in[s,T]} \sup_{z\in\mathbb{R}^d} \frac{\mathbb{E}\Big[\big|U^0_{n-1,M}(r,z) - u(r,z)\big|^2\Big]}{\varphi(z)}\right] e^{\rho(s-t)}\varphi(x)\, ds\right)^{1/2}.
\end{aligned} \tag{79}$$

Therefore, we obtain that

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d}\left[\frac{e^{\rho t}\,\big|\mathbb{E}\big[U^0_{n,M}(t,x)\big] - u(t,x)\big|}{\varphi(x)}\right]^{1/2} \\
&\leq L(T-t)^{1/2}\left(\int_t^T \sup_{r\in[s,T]} \sup_{x\in\mathbb{R}^d} \left[\frac{e^{\rho r}\mathbb{E}\big[|U^0_{n-1,M}(r,x) - u(r,x)|^2\big]}{\varphi(x)}\right] ds\right)^{1/2}.
\end{aligned} \tag{80}$$

Combining (77) and the triangle inequality hence shows that

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}^d}\Big[e^{\rho t}\varphi(x)^{-1}\mathbb{E}\Big[\big|U^0_{n,M}(t,x)-u(t,x)\big|^2\Big]\Big]^{1/2}\\
&\le \sup_{x\in\mathbb{R}^d}\Big[e^{\rho t}\varphi(x)^{-1}\left|\mathbb{E}\big[U^0_{n,M}(t,x)\big]-u(t,x)\right|\Big]^{1/2}+\sup_{x\in\mathbb{R}^d}\Big[e^{\rho t}\varphi(x)^{-1}\mathbb{V}\mathrm{ar}\big(U^0_{n,M}(t,x)\big)\Big]^{1/2}\\
&\le L(T-t)^{1/2}\left(\int_t^T\sup_{r\in[s,T]}\sup_{x\in\mathbb{R}^d}\Big[(\varphi(x))^{-1}e^{\rho r}\mathbb{E}\Big[\big|U^0_{n-1,M}(r,x)-u(r,x)\big|^2\Big]\Big]\,ds\right)^{1/2}\\
&\quad+\frac{2e^{\rho T/2}}{\sqrt{M^n}}\sup_{s\in[0,T]}\sup_{z\in\mathbb{R}^d}\left[\frac{\max\{|T(F(0))(s,z)|,|g(z)|\}}{\sqrt{\varphi(z)}}\right]\\
&\quad+\left[\sum_{\ell=0}^{n-1}\frac{(2-\mathbb{1}_{\{n-1\}}(\ell))(T-t)^{1/2}L}{\sqrt{M^{n-\ell-1}}}\left(\int_t^T\sup_{r\in[s,T]}\sup_{x\in\mathbb{R}^d}\left[\frac{e^{\rho r}\mathbb{E}\big[|U^0_{\ell,M}(r,x)-u(r,x)|^2\big]}{\varphi(x)}\right]ds\right)^{1/2}\right]\\
&=\frac{2e^{\rho T/2}}{\sqrt{M^n}}\sup_{s\in[0,T]}\sup_{z\in\mathbb{R}^d}\left[\frac{\max\{|T(F(0))(s,z)|,|g(z)|\}}{\sqrt{\varphi(z)}}\right]\\
&\quad+\sum_{\ell=0}^{n-1}\frac{2(T-t)^{1/2}L}{\sqrt{M^{n-\ell-1}}}\left(\int_t^T\sup_{r\in[s,T]}\sup_{x\in\mathbb{R}^d}\left[\frac{e^{\rho r}\mathbb{E}\big[|U^0_{\ell,M}(r,x)-u(r,x)|^2\big]}{\varphi(x)}\right]ds\right)^{1/2}.
\end{aligned}
\tag{81}
$$

This establishes (72). The proof of Lemma 3.5 is thus complete. □

3.4 Function space-valued Gronwall-type inequalities

Lemma 3.6. *Let $K\in\mathbb{N}$, $\alpha,\beta\in[0,\infty)$, $\varepsilon_0,\varepsilon_1,\dots\varepsilon_K\in[0,\infty]$ satisfy $\max\{\varepsilon_0,\varepsilon_1,\dots\varepsilon_{K-1}\}<\infty$ and $\varepsilon_K\le\alpha+\beta\big[\sum_{k=0}^{K-1}\varepsilon_k\big]$. Then $\varepsilon_K<\infty$.*

Proof of Lemma 3.6. Note that the hypothesis that $\max\{\varepsilon_0,\varepsilon_1,\dots\varepsilon_{K-1}\}<\infty$ implies that $\alpha+\beta\big[\sum_{k=0}^{K-1}\varepsilon_k\big]<\infty$. This and the hypothesis that $\varepsilon_K\le\alpha+\beta\big[\sum_{k=0}^{K-1}\varepsilon_k\big]$ establish that $\varepsilon_K<\infty$. The proof of Lemma 3.6 is thus complete. □

Lemma 3.7. *Let $N\in\mathbb{N}$, $\beta,\alpha_0,\alpha_1,\dots,\alpha_N\in[0,\infty)$, $\varepsilon_0,\varepsilon_1,\dots\varepsilon_N\in[0,\infty]$ satisfy for all $n\in\{0,1,\dots,N\}$ that $\varepsilon_n\le\alpha_n+\beta\big[\sum_{k=0}^{n-1}\varepsilon_k\big]$ (cf. Lemma 3.6). Then it holds for all $n\in\{0,1,\dots,N\}$ that*

$$\varepsilon_n\le\alpha_n+\beta\left[\sum_{k=0}^{n-1}(1+\beta)^{n-k-1}\alpha_k\right]<\infty. \tag{82}$$

Proof of Lemma 3.7. Throughout this proof let $\gamma_0,\gamma_1,\dots,\gamma_N\in\mathbb{R}$ satisfy for all $n\in\{0,1,\dots,N\}$ that

$$\gamma_n=\alpha_n+\beta\left[\sum_{k=0}^{n-1}\gamma_k\right]. \tag{83}$$

We claim that for all $n\in\{0,1,\dots,N\}$ it holds that

$$\gamma_n=\alpha_n+\beta\left[\sum_{k=0}^{n-1}(1+\beta)^{n-k-1}\alpha_k\right]. \tag{84}$$

We prove (84) by induction on $n\in\{0,1,\dots,N\}$. For the base case $n=0$ observe that (83) ensures that $\gamma_0=\alpha_0$. This proves (84) in the base case $n=0$. For the induction step $\{0,1,\dots,N-1\} \ni n-1 \dashrightarrow n \in \{1,2,\dots,N\}$ observe that (83) implies that for all $n \in \{1,2,\dots,N\}$ with $\gamma_{n-1} = \alpha_{n-1} + \beta \sum_{k=0}^{n-2}(1+\beta)^{n-k-2}\alpha_k$ it holds that

$$\begin{aligned}\gamma_n &= \alpha_n + \beta\left[\sum_{k=0}^{n-1}\gamma_k\right] = \alpha_n - \alpha_{n-1} + \beta\gamma_{n-1} + \alpha_{n-1} + \beta\left[\sum_{k=0}^{n-2}\gamma_k\right] \\ &= \alpha_n - \alpha_{n-1} + \beta\gamma_{n-1} + \gamma_{n-1} = \alpha_n - \alpha_{n-1} + (1+\beta)\gamma_{n-1} \\ &= \alpha_n - \alpha_{n-1} + (1+\beta)\left(\alpha_{n-1} + \beta\left[\sum_{k=0}^{n-2}(1+\beta)^{n-k-2}\alpha_k\right]\right) \\ &= \alpha_n + \beta\alpha_{n-1} + \beta\left[\sum_{k=0}^{n-2}(1+\beta)^{n-k-1}\alpha_k\right] = \alpha_n + \beta\left[\sum_{k=0}^{n-1}(1+\beta)^{n-k-1}\alpha_k\right].\end{aligned} \tag{85}$$

Induction hence establishes (84). Moreover, note that (83), induction, and the assumption that for all $n \in \{0,1,\dots,N\}$ it holds that $\varepsilon_n \le \alpha_n + \beta\big[\sum_{k=0}^{n-1}\varepsilon_k\big]$ prove that for all $n \in \{0,1,\dots,N\}$ it holds that $\varepsilon_n \le \gamma_n$. This and (84) establish that for all $n \in \{0,1,\dots,N\}$ it holds that

$$\varepsilon_n \le \alpha_n + \beta\left[\sum_{k=0}^{n-1}(1+\beta)^{n-k-1}\alpha_k\right] < \infty. \tag{86}$$

The proof of Lemma 3.7 is thus complete. $\square$

Lemma 3.8. *Let $K \in \mathbb{N}$, $a,b,c \in [0,\infty)$, $\alpha \in \mathbb{R}$, $\beta \in [\alpha,\infty)$, $p \in (0,\infty)$, let $f_n\colon [\alpha,\beta] \to [0,\infty]$, $n \in \mathbb{N}_0$, be measurable, assume $\sup_{s\in[\alpha,\beta]}\max\{|f_0(s)|,|f_1(s)|,\dots,|f_{K-1}(s)|\} < \infty$, and assume for all $t \in [\alpha,\beta]$ that*

$$|f_K(t)| \le ac^K + \sum_{\ell=0}^{K-1}\left[bc^{K-\ell-1}\left[\int_\alpha^t |f_\ell(s)|^p\,ds\right]^{1/p}\right]. \tag{87}$$

Then $\sup_{s\in[\alpha,\beta]}|f_K(s)| < \infty$.

Proof of Lemma 3.8. Note that the hypothesis that $\sup_{s\in[\alpha,\beta]}\max\{|f_0(s)|,|f_1(s)|,\dots,|f_{K-1}(s)|\} < \infty$ implies that

$$\begin{aligned}&\sup_{t\in[\alpha,\beta]}\left(\sum_{\ell=0}^{K-1}\left[bc^{K-\ell-1}\left[\int_\alpha^t |f_\ell(s)|^p\,ds\right]^{1/p}\right]\right) \le \sum_{\ell=0}^{K-1}\left[bc^{K-\ell-1}\left[\int_\alpha^\beta |f_\ell(s)|^p\,ds\right]^{1/p}\right] \\ &\le \sum_{\ell=0}^{K-1}\left[bc^{K-\ell-1}\left[\sup_{s\in[\alpha,\beta]}|f_\ell(s)|\right][\beta-\alpha]^{1/p}\right] < \infty.\end{aligned} \tag{88}$$

Combining this with (87) establishes that $\sup_{s\in[\alpha,\beta]}|f_K(s)| < \infty$. The proof of Lemma 3.8 is thus complete. $\square$

Lemma 3.9. *Let $N \in \mathbb{N}$, $a,b,c \in [0,\infty)$, $\alpha \in \mathbb{R}$, $\beta \in [\alpha,\infty)$, $p \in [1,\infty)$, let $f_n\colon [\alpha,\beta] \to [0,\infty]$, $n \in \mathbb{N}_0$, be measurable, assume $\sup_{s\in[\alpha,\beta]}|f_0(s)| < \infty$, and assume for all $n \in \{1,2,\dots,N\}$, $t \in [\alpha,\beta]$ that*

$$|f_n(t)| \le ac^n + \sum_{\ell=0}^{n-1}\left[bc^{n-\ell-1}\left[\int_\alpha^t |f_\ell(s)|^p\,ds\right]^{1/p}\right] \tag{89}$$

(cf. Lemma 3.8). Then

$$\begin{aligned} f_N(\beta) &\leq ac^N + b(\beta-\alpha)^{1/p}[1+b(\beta-\alpha)^{1/p}]^{N-1}\left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}}\right]\left[\sup_{s\in[\alpha,\beta]}|f_0(s)|\right] \\ &\quad + ab(\beta-\alpha)^{1/p}\sum_{n=1}^{N-1}[1+b(\beta-\alpha)^{1/p}]^{N-n-1}\left[\max_{k\in\{0,1,\dots,N-n\}} \tfrac{c^{N-k}}{(k!)^{1/p}}\right] \\ &\leq \left[a + b(\beta-\alpha)^{1/p}\left[\sup_{s\in[\alpha,\beta]}|f_0(s)|\right]\right]\left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}}\right]\left[1+b(\beta-\alpha)^{1/p}\right]^{N-1}. \end{aligned} \tag{90}$$

Proof of Lemma 3.9. Throughout this proof assume w.l.o.g. that $\alpha < \beta$, let $\gamma_k \colon \mathcal{B}([\alpha,\beta]) \to [0,\infty)$, $k \in \mathbb{N}_0$, satisfy for all $k \in \mathbb{N}$, $A \in \mathcal{B}([\alpha,\beta])$ that

$$\gamma_0(A) = \mathbb{1}_A(\beta) \qquad \text{and} \qquad \gamma_k(A) = \frac{1}{(\beta-\alpha)^k}\int_A \frac{(\beta-t)^{k-1}}{(k-1)!}\,dt, \tag{91}$$

and let $\varepsilon_n \in [0,\infty]$, $n \in \mathbb{N}_0$, satisfy for all $n \in \mathbb{N}_0$ that

$$\varepsilon_n = \sup\left\{c^j\left[\int_\alpha^\beta |f_n(t)|^p\gamma_k(dt)\right]^{1/p} : j,k\in\mathbb{N}_0, n+j+k=N\right\}. \tag{92}$$

Observe that (91) ensures that for all $k \in \mathbb{N}_0$ it holds that

$$\int_\alpha^\beta \gamma_k(dt) = \frac{1}{k!} \tag{93}$$

This and (92) show that

$$\begin{aligned} \varepsilon_0 &= \sup\left\{c^j\left[\int_\alpha^\beta |f_0(t)|^p\gamma_k(dt)\right]^{1/p} : j,k\in\mathbb{N}_0, j+k=N\right\} \\ &\leq \sup\left\{c^j|f_0(s)|\left[\int_\alpha^\beta \gamma_k(dt)\right]^{1/p} : j,k\in\mathbb{N}_0, j+k=N, s\in[\alpha,\beta]\right\} \\ &= \sup\left\{\frac{c^j|f_0(s)|}{(k!)^{1/p}} \colon j,k\in\mathbb{N}_0, j+k=N, s\in[\alpha,\beta]\right\} \leq \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}}\right]\left[\sup_{s\in[\alpha,\beta]}|f_0(s)|\right]. \end{aligned} \tag{94}$$

Next note that Fubini's theorem and (91) imply that for all $\ell \in [0,N]\cap\mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned} &\int_\alpha^\beta\int_\alpha^t |f_\ell(s)|^p\,ds\,\gamma_k(dt) \\ &= \mathbb{1}_{\{0\}}(k)\left[\int_\alpha^\beta |f_\ell(s)|^p\,ds\right] + \mathbb{1}_{\mathbb{N}}(k)\left[\frac{1}{(\beta-\alpha)^k}\int_\alpha^\beta \frac{(\beta-t)^{k-1}}{(k-1)!}\int_\alpha^t |f_\ell(s)|^p\,ds\,dt\right] \\ &= \mathbb{1}_{\{0\}}(k)\left[\int_\alpha^\beta |f_\ell(s)|^p\,ds\right] + \mathbb{1}_{\mathbb{N}}(k)\left[\frac{1}{(\beta-\alpha)^k}\int_\alpha^\beta\int_s^\beta \frac{(\beta-t)^{k-1}}{(k-1)!}|f_\ell(s)|^p\,dt\,ds\right] \\ &= \mathbb{1}_{\{0\}}(k)\left[\int_\alpha^\beta |f_\ell(s)|^p\,ds\right] + \mathbb{1}_{\mathbb{N}}(k)\left[\frac{1}{(\beta-\alpha)^k}\int_\alpha^\beta \frac{(\beta-s)^k}{k!}|f_\ell(s)|^p\,ds\right] \\ &= \mathbb{1}_{\{0\}}(k)\left[(\beta-\alpha)\int_\alpha^\beta |f_\ell(s)|^p\,\gamma_1(ds)\right] + \mathbb{1}_{\mathbb{N}}(k)\left[(\beta-\alpha)\int_\alpha^\beta |f_\ell(s)|^p\,\gamma_{k+1}(ds)\right] \\ &= (\beta-\alpha)\int_\alpha^\beta |f_\ell(s)|^p\,\gamma_{k+1}(ds). \end{aligned} \tag{95}$$

Combining (89) and (93) hence assures that for all $n \in \{1, 2, \dots, N\}$, $k, j \in \mathbb{N}_0$ with $n+k+j = N$ it holds that

$$\begin{aligned} & c^j \left[\int_\alpha^\beta |f_n(t)|^p \gamma_k(dt) \right]^{1/p} \\ & \leq a c^{n+j} \left[\int_\alpha^\beta \gamma_k(dt) \right]^{1/p} + \sum_{\ell=0}^{n-1} \left[b c^{n+j-\ell-1} \left[\int_\alpha^\beta \int_\alpha^t |f_\ell(s)|^p \, ds \, \gamma_k(dt) \right]^{1/p} \right] \\ & \leq \frac{a c^{N-k}}{(k!)^{1/p}} + \sum_{\ell=0}^{n-1} \left[b(\beta-\alpha)^{1/p} c^{n+j-\ell-1} \left[\int_\alpha^\beta |f_\ell(s)|^p \gamma_{k+1}(ds) \right]^{1/p} \right]. \end{aligned} \tag{96}$$

This, (92), and the fact that for all $n, k, j \in \mathbb{N}_0$ with $n + k + j = N$ it holds that $\ell + (n + j - \ell - 1) + (k + 1) = N$ imply that for all $n \in \{1, 2, \dots, N\}$ it holds that

$$\varepsilon_n \leq a \left[\max_{k \in \{0,1,\dots,N-n\}} \frac{c^{N-k}}{(k!)^{1/p}} \right] + b(\beta-\alpha)^{1/p} \left[\sum_{\ell=0}^{n-1} \varepsilon_\ell \right]. \tag{97}$$

Combining (91), (92), and (94) with Lemma 3.7 (applied with $\beta \curvearrowleft b(\beta-\alpha)^{1/p}$, $\alpha_0 \curvearrowleft [\max_{k\in\{0,1,\dots,N\}} \frac{c^{N-k}}{(k!)^{1/p}}][\sup_{s\in[\alpha,\beta]} |f_0(s)|]$, $(\alpha_n)_{n\in\{1,2,\dots,N\}} \curvearrowleft (a[\max_{k\in\{0,1,\dots,N-n\}} \frac{c^{N-k}}{(k!)^{1/p}}])_{n\in\{1,2,\dots,N\}}$ in the notation of Lemma 3.7) hence shows that

$$\begin{aligned} f_N(\beta) &= \left[\int_\alpha^\beta |f_N(t)|^p \, \gamma_0(dt) \right]^{1/p} = \varepsilon_N \\ & \leq a c^N + b(\beta-\alpha)^{1/p} [1 + b(\beta-\alpha)^{1/p}]^{N-1} \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] \left[\sup_{s\in[\alpha,\beta]} |f_0(s)| \right] \\ & + ab(\beta-\alpha)^{1/p} \sum_{n=1}^{N-1} [1 + b(\beta-\alpha)^{1/p}]^{N-n-1} \left[\max_{k\in\{0,1,\dots,N-n\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right]. \end{aligned} \tag{98}$$

The fact that for all $l \in \{0, 1, \dots, N\}$ it holds that $\max_{k\in\{0,1,\dots,l\}} \frac{c^{N-k}}{(k!)^{1/p}} \leq \max_{k\in\{0,1,\dots,N\}} \frac{c^{N-k}}{(k!)^{1/p}}$ therefore implies that

$$\begin{aligned} \varepsilon_N & \leq a \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] + ab(\beta-\alpha)^{1/p} \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] \sum_{n=1}^{N-1} [1 + b(\beta-\alpha)^{1/p}]^{N-n-1} \\ & + b(\beta-\alpha)^{1/p} [1 + b(\beta-\alpha)^{1/p}]^{N-1} \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] \left[\sup_{s\in[\alpha,\beta]} |f_0(s)| \right] \\ & = a \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] [1 + b(\beta-\alpha)^{1/p}]^{N-1} \\ & + b(\beta-\alpha)^{1/p} [1 + b(\beta-\alpha)^{1/p}]^{N-1} \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] \left[\sup_{s\in[\alpha,\beta]} |f_0(s)| \right] \\ & = \left[\max_{k\in\{0,1,\dots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] [1 + b(\beta-\alpha)^{1/p}]^{N-1} \left[a + b(\beta-\alpha)^{1/p} \left[\sup_{s\in[\alpha,\beta]} |f_0(s)| \right] \right]. \end{aligned} \tag{99}$$

Combining this with (98) establishes (90). The proof of Lemma 3.9 is thus complete. ☐

Lemma 3.10. *Let $N \in \mathbb{N}$, $a, b, c \in [0, \infty)$, $\alpha \in \mathbb{R}$, $\beta \in [\alpha, \infty)$, $p \in [1, \infty)$, let $f_n \colon [\alpha, \beta] \to [0, \infty]$, $n \in \mathbb{N}_0$, be measurable, assume $\sup_{s\in[\alpha,\beta]} |f_0(s)| < \infty$, and assume for all $n \in \{1, 2, \dots, N\}$, $t \in [\alpha, \beta]$ that*

$$|f_n(t)| \leq a c^n + \sum_{\ell=0}^{n-1} \left[b c^{n-\ell-1} \left[\int_t^\beta |f_\ell(s)|^p \, ds \right]^{1/p} \right] \tag{100}$$

(cf. Lemma 3.8). Then

$$f_N(\alpha) \le \left[a + b(\beta-\alpha)^{1/p} \left[\sup_{s\in[\alpha,\beta]} |f_0(s)| \right] \right] \left[\max_{k\in\{0,1,\ldots,N\}} \tfrac{c^{N-k}}{(k!)^{1/p}} \right] \left[1 + b(\beta-\alpha)^{1/p} \right]^{N-1}. \tag{101}$$

Proof of Lemma 3.10. Note that Lemma 3.9 (applied with $\alpha \curvearrowleft -\beta$, $\beta \curvearrowleft -\alpha$, $(f_n)_{n\in\mathbb{N}_0} \curvearrowleft (([-\beta,-\alpha] \ni t \mapsto f_n(-t) \in [0,\infty]))_{n\in\mathbb{N}_0}$ in the notation of Lemma 3.9) establishes (101). The proof of Lemma 3.10 is thus complete. ◻

Lemma 3.11. *Let $M, N \in \mathbb{N}$, $T \in (0,\infty)$, $\tau \in [0,T]$, $a, b \in [0,\infty)$, $p \in [1,\infty)$, let $f_n \colon [\tau,T] \to [0,\infty]$, $n \in \mathbb{N}_0$, be measurable, assume $\sup_{s\in[\tau,T]} |f_0(s)| < \infty$, and assume for all $n \in \{1,2,\ldots,N\}$, $t \in [\tau,T]$ that*

$$|f_n(t)| \le \frac{a}{\sqrt{M^n}} + \sum_{\ell=0}^{n-1} \left[\frac{b}{\sqrt{M^{n-\ell-1}}} \left[\int_t^T |f_\ell(s)|^p \, ds \right]^{1/p} \right] \tag{102}$$

(cf. Lemma 3.8). Then

$$f_N(\tau) \le \left[a + b(T-\tau)^{1/p} \left[\sup_{s\in[\tau,T]} |f_0(s)| \right] \right] \exp\!\left(\frac{M^{p/2}}{p} \right) M^{-N/2} \left[1 + b(T-\tau)^{1/p} \right]^{N-1}. \tag{103}$$

Proof of Lemma 3.11. Note that Lemma 3.10 (applied with $c \curvearrowleft M^{-1/2}$, $\alpha \curvearrowleft \tau$, $\beta \curvearrowleft T$ in the notation of Lemma 3.10) assures that

$$f_N(\tau) \le \left[a + b(T-\tau)^{1/p} \left[\sup_{s\in[\tau,T]} |f_0(s)| \right] \right] \left[\sup_{k\in\mathbb{N}_0} \tfrac{M^{k/2}}{(k!)^{1/p}} \right] M^{-N/2} \left[1 + b(T-\tau)^{1/p} \right]^{N-1}. \tag{104}$$

The fact that $\sup_{k\in\mathbb{N}_0}\big(\frac{M^{k/2}}{(k!)^{1/p}}\big) \le \exp\big(\frac{M^{p/2}}{p}\big)$ hence proves that

$$f_N(\tau) \le \left[a + b(T-\tau)^{1/p} \left[\sup_{s\in[\tau,T]} |f_0(s)| \right] \right] \exp\!\left(\frac{M^{p/2}}{p} \right) M^{-N/2} \left[1 + b(T-\tau)^{1/p} \right]^{N-1}. \tag{105}$$

The proof of Lemma 3.11 is thus complete. ◻

3.5 Non-recursive error bounds for MLP approximations

Corollary 3.12 (Error estimate)**.** *Assume Setting 3.1 and let $N, M \in \mathbb{N}$, $\tau \in [0,T]$. Then*

$$\begin{aligned} &\sup_{x\in\mathbb{R}^d} \left[\frac{\mathbb{E}\big[|U^0_{N,M}(\tau,x) - u(\tau,x)|^2\big]}{\varphi(x)} \right]^{1/2} \le 2e^{M/2} M^{-N/2} (1+2TL)^{N-1} e^{\rho(T-\tau)/2} \\ &\cdot \left[\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[\frac{\max\{|T(F(0))(t,x)|, |g(x)|\}}{\sqrt{\varphi(x)}} \right] + TL \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left[\frac{|u(t,x)|}{\sqrt{\varphi(x)}} \right] \right]. \end{aligned} \tag{106}$$

Proof of Corollary 3.12. Throughout this proof let $a_1, a_2 \in \mathbb{R}$ satisfy

$$a_1 = 2e^{\rho T/2} \left[\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \left(\frac{\max\{|T(F(0))(t,x)|, |g(x)|\}}{\sqrt{\varphi(x)}} \right) \right] \quad \text{and} \quad a_2 = 2\sqrt{T}L, \tag{107}$$

let $f_n \colon [\tau,T] \to [0,\infty]$, $n \in \{0,1,\ldots,N\}$, satisfy for all $n \in \{0,1,\ldots,N\}$, $t \in [\tau,T]$ that

$$f_n(t) = \sup_{s\in[t,T]} \sup_{x\in\mathbb{R}^d} \left[e^{\rho s} |\varphi(x)|^{-1} \mathbb{E}\big[|U^0_{n,M}(s,x) - u(s,x)|^2\big] \right]^{1/2}. \tag{108}$$

Observe that (108) ensures that for all $n \in \{0, 1, \ldots, N\}$ it holds that f_n is measurable. Furthermore, note that (108) and Lemma 3.5 imply that for all $n \in \{1, 2, \ldots, N\}$, $t \in [\tau, T]$ it holds that

$$\begin{aligned} |f_n(t)| &\leq \sup_{r \in [t,T]} \left[\frac{a_1}{\sqrt{M^n}} + \left[\sum_{\ell=0}^{n-1} \frac{a_2}{\sqrt{M^{n-\ell-1}}} \left[\int_r^T |f_\ell(s)|^2 \, ds \right]^{1/2} \right] \right] \\ &\leq \frac{a_1}{\sqrt{M^n}} + \left[\sum_{\ell=0}^{n-1} \frac{a_2}{\sqrt{M^{n-\ell-1}}} \left[\int_t^T |f_\ell(s)|^2 \, ds \right]^{1/2} \right]. \end{aligned} \tag{109}$$

Lemma 3.11, (108), and the fact that for all $t \in [\tau, T]$ it holds that

$$f_0(t) = \sup_{s \in [t,T]} \sup_{x \in \mathbb{R}^d} \left(\frac{e^{\rho s/2} |u(s,x)|}{\sqrt{\varphi(x)}} \right) \leq e^{\rho T/2} \left[\sup_{s \in [0,T]} \sup_{x \in \mathbb{R}^d} \left(\frac{|u(s,x)|}{\sqrt{\varphi(x)}} \right] \right) < \infty \tag{110}$$

therefore demonstrate that

$$\begin{aligned} &\sup_{t \in [\tau,T]} \sup_{x \in \mathbb{R}^d} \left[e^{\rho t} \varphi(x)^{-1} \mathbb{E}\left[\left| U^0_{N,M}(t,x) - u(t,x) \right|^2 \right] \right]^{1/2} = f_N(\tau) \\ &\leq \left[a_1 + a_2 \sqrt{T} \sup\nolimits_{t \in [\tau,T)} |f_0(t)| \right] e^{M/2} M^{-N/2} (1 + a_2 \sqrt{T})^{N-1} \\ &\leq e^{M/2} M^{-N/2} (1 + 2TL)^{N-1} \\ &\cdot \left[2 e^{\rho T/2} \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left[\frac{\max\{|T(F(0))(t,x)|, |g(x)|\}}{\sqrt{\varphi(x)}} \right] + 2TL e^{\rho T/2} \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left[\frac{|u(t,x)|}{\sqrt{\varphi(x)}} \right] \right]. \end{aligned} \tag{111}$$

This completes the proof of Corollary 3.12. □

4 Computational complexity analysis for MLP approximations

In this section we combine the existence, uniqueness, and regularity properties for solutions of stochastic fixed point equations, which we have established in Section 2, with the error analysis for MLP approximations for stochastic fixed point equations, which we have established in Section 3 (see Corollary 3.12 in Subsection 3.5), to obtain in Theorem 4.2 in Subsection 4.2 a computational complexity analysis for MLP approximations for semilinear second-order PDEs in fixed space dimensions. In Subsection 4.3 we combine the computational complexity analysis in Theorem 4.2 with the elementary auxiliary result in Lemma 4.3 to obtain in Corollary 4.4 a computational complexity analysis for MLP approximations for semilinear second-order PDEs in variable space dimensions.

4.1 Error bounds for MLP approximations involving Euler-Maruyama approximations

Proposition 4.1. *Let $d, m, M, K \in \mathbb{N}$, $\beta, b, c \in [1, \infty)$, $p \in [2\beta, \infty)$, $\varphi \in C^2(\mathbb{R}^d, [1, \infty))$, $g \in C(\mathbb{R}^d, \mathbb{R})$, $\mu \in C(\mathbb{R}^d, \mathbb{R}^d)$, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_m) \in C(\mathbb{R}^d, \mathbb{R}^{d \times m})$, $T, \tau_0, \tau_1, \ldots, \tau_K \in \mathbb{R}$ satisfy $0 = \tau_0 < \tau_1 < \ldots < \tau_K = T$, let $\|\cdot\| \colon \mathbb{R}^d \to [0, \infty)$ be the standard norm on $\mathbb{R}^d$, let $f \colon [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ be measurable, let $F \colon \mathbb{R}^{[0,T] \times \mathbb{R}^d} \to \mathbb{R}^{[0,T] \times \mathbb{R}^d}$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v \in \mathbb{R}^{[0,T] \times \mathbb{R}^d}$ that $(F(v))(t,x) = f(t, x, v(t,x))$, assume for all $x, y \in \mathbb{R}^d$, $z \in \mathbb{R}^d \backslash \{0\}$, $t \in [0,T]$, $v, w \in \mathbb{R}$ that*

$$\max\left\{ \tfrac{|(\varphi'(x))(z)|}{(\varphi(x))^{(p-1)/p} \|z\|}, \tfrac{|(\varphi''(x))(z,z)|}{(\varphi(x))^{(p-2)/p} \|z\|^2}, \tfrac{c\|x\| + \|\mu(0)\|}{(\varphi(x))^{1/p}}, \tfrac{c\|x\| + [\sum_{i=1}^m \|\sigma_i(0)\|^2]^{1/2}}{(\varphi(x))^{1/p}} \right\} \leq c, \tag{112}$$

$$\max\{|Tf(t,x,0)|,|g(x)|\} \leq b(\varphi(x))^{\beta/p}, \tag{113}$$

$$\max\{|g(x)-g(y)|,T|f(t,x,v)-f(t,y,w)|\} \leq cT|v-w| + \tfrac{(\varphi(x)+\varphi(y))^{\beta/p}\|x-y\|}{T^{1/2}b^{-1}}, \tag{114}$$

$$\max\{\|\mu(x)-\mu(y)\|^2, \textstyle\sum_{i=1}^m \|\sigma_i(x)-\sigma_i(y)\|^2\} \leq c^2\|x-y\|^2, \tag{115}$$

let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ *be a filtered probability space which satisfies the usual conditions*[4], *let* $\Theta = \bigcup_{n\in\mathbb{N}} \mathbb{Z}^n$, *let* $\mathfrak{r}^\theta \colon \Omega \to [0,1]$, $\theta \in \Theta$, *be i.i.d. random variables, assume for all* $t \in (0,1)$ *that* $\mathbb{P}(\mathfrak{r}^0 \leq t) = t$, *let* $\mathcal{R}^\theta \colon [0,T]\times\Omega \to [0,T]$, $\theta\in\Theta$, *satisfy for all* $\theta\in\Theta$, $t\in[0,T]$ *that* $\mathcal{R}^\theta_t = t + (T-t)\mathfrak{r}^\theta$, *let* $W^\theta \colon [0,T]\times\Omega \to \mathbb{R}^m$, $\theta\in\Theta$, *be i.i.d. standard* $(\mathbb{F}_t)_{t\in[0,T]}$*-Brownian motions, assume that* $(\mathfrak{r}^\theta)_{\theta\in\Theta}$ *and* $(W^\theta)_{\theta\in\Theta}$ *are independent, let* $\llcorner\cdot\lrcorner \colon \mathbb{R}\to\mathbb{R}$ *satisfy for all* $t\in\mathbb{R}$ *that* $\llcorner t \lrcorner = \max(\{\tau_0,\tau_1,\ldots,\tau_n\}\cap((-\infty,t)\cup\{\tau_0\}))$, *for every* $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *let* $Y^{\theta,x}_t = (Y^{\theta,x}_{t,s})_{s\in[t,T]} \colon [t,T]\times\Omega\to\mathbb{R}^d$ *satisfy for all* $s\in[t,T]$ *that* $Y^{\theta,x}_{t,t} = x$ *and*

$$\begin{aligned} &Y^{\theta,x}_{t,s} - Y^{\theta,x}_{t,\max\{t,\llcorner s\lrcorner\}} \\ &= \mu(Y^{\theta,x}_{t,\max\{t,\llcorner s\lrcorner\}})(s-\max\{t,\llcorner s\lrcorner\}) + \sigma(Y^{\theta,x}_{t,\max\{t,\llcorner s\lrcorner\}})(W^\theta_s - W^\theta_{\max\{t,\llcorner s\lrcorner\}}), \end{aligned} \tag{116}$$

and let $U^\theta_n \colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $n\in\mathbb{Z}$, $\theta\in\Theta$, *satisfy for all* $\theta\in\Theta$, $n\in\mathbb{N}_0$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *that*

$$\begin{aligned} U^\theta_n(t,x) &= \frac{\mathbb{1}_{\mathbb{N}}(n)}{M^n}\sum_{i=1}^{M^n} g\big(Y^{(\theta,0,-i),x}_{t,T}\big) \\ &+ \sum_{\ell=0}^{n-1}\frac{(T-t)}{M^{n-\ell}}\left[\sum_{i=1}^{M^{n-\ell}}\big(F\big(U^{(\theta,\ell,i)}_\ell\big) - \mathbb{1}_{\mathbb{N}}(\ell)F\big(U^{(\theta,-\ell,i)}_{\ell-1}\big)\big)\big(\mathcal{R}^{(\theta,\ell,i)}_t, Y^{(\theta,\ell,i),x}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}\big)\right]. \end{aligned} \tag{117}$$

Then

(i) for every $t\in[0,T]$, $\theta\in\Theta$ *there exists an up to indistinguishability unique continuous random field* $(X^{\theta,x}_{t,s})_{(s,x)\in[t,T]\times\mathbb{R}^d} \colon [t,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ *which satisfies that for all* $x\in\mathbb{R}^d$ *it holds that* $(X^{\theta,x}_{t,s})_{s\in[t,T]}$ *is* $(\mathbb{F}_s)_{s\in[t,T]}$*-adapted and which satisfies that for all* $s\in[t,T]$, $x\in\mathbb{R}^d$ *it holds* $\mathbb{P}$*-a.s. that*

$$X^{\theta,x}_{t,s} = x + \int_t^s \mu(X^{\theta,x}_{t,r})\,dr + \int_t^s \sigma(X^{\theta,x}_{t,r})\,dW^\theta_r, \tag{118}$$

(ii) it holds for all $\theta\in\Theta$, $t\in[0,T]$, $s\in[t,T]$, $r\in[s,T]$, $x\in\mathbb{R}^d$ *that* $\mathbb{P}\big(X^{\theta,X^{\theta,x}_{t,s}}_{s,r} = X^{\theta,x}_{t,r}\big) = 1$,

(iii) there exists a unique measurable $u\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$ *which satisfies for all* $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $\big(\sup_{s\in[0,T],y\in\mathbb{R}^d}[|u(s,y)|(\varphi(y))^{-\beta/p}]\big) + \int_t^T \mathbb{E}\big[|f(s,X^{0,x}_{t,s},u(s,X^{0,x}_{t,s}))|\big]\,ds + \mathbb{E}\big[|g(X^{0,x}_{t,T})|\big] < \infty$ *and*

$$u(t,x) = \mathbb{E}\big[g(X^{0,x}_{t,T})\big] + \int_t^T \mathbb{E}\big[f(s,X^{0,x}_{t,s},u(s,X^{0,x}_{t,s}))\big]\,ds, \tag{119}$$

(iv) it holds for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $n\in\mathbb{N}_0$, $\theta\in\Theta$ *that* $U^\theta_n(t,x)$ *is measurable, and*

(v) it holds for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $n\in\mathbb{N}_0$ *that*

$$\frac{\big(\mathbb{E}\big[|U^0_n(t,x)-u(t,x)|^2\big]\big)^{1/2}}{12bc^2\exp(9c^3T)|\varphi(x)|^{(\beta+1)/p}} \leq \left[\frac{\exp(2ncT+\frac{M}{2})}{M^{n/2}} + \max_{i\in\{1,2,\ldots,K\}}\Big(\tfrac{|\tau_i-\tau_{i-1}|^{1/2}}{T^{1/2}}\Big)\right]. \tag{120}$$

[4]Let $T\in[0,\infty)$ and let $\mathbf{\Omega} = (\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space. Then we say that $\mathbf{\Omega}$ satisfies the usual conditions if and only if it holds that $\{A\in\mathcal{F}: \mathbb{P}(A)=0\}\subseteq\mathbb{F}_0$ and $\forall\, t\in[0,T)\colon \mathbb{F}_t = \cap_{s\in(t,T]}\mathbb{F}_s$.

Proof of Proposition 4.1. Throughout this proof let $|||\cdot|||\colon (\bigcup_{L,N\in\mathbb{N}}\mathbb{R}^{L\times N})\to[0,\infty)$ satisfy for all $L,N\in\mathbb{N}$, $A=(A_{i,j})_{(i,j)\in\{1,2,\ldots,L\}\times\{1,2,\ldots,N\}}\in\mathbb{R}^{L\times N}$ that $|||A|||=[\sum_{i=1}^L\sum_{j=1}^N|A_{ij}|^2]^{1/2}$, let $\delta=\max_{i\in\{1,2,\ldots,K\}}|\tau_i-\tau_{i-1}|$, let $\Delta\subseteq[0,T]^2$ satisfy $\Delta=\{(t,s)\in[0,T]^2\colon t\leq s\}$, let $\mathfrak{X}^k=(\mathfrak{X}^{k,x}_{t,s})_{(t,s,x)\in\Delta\times\mathbb{R}^d}\colon\Delta\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$, $k\in\{0,1\}$, satisfy for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ that $\mathfrak{X}^{0,x}_{t,s}=X^{0,x}_{t,s}$ and $\mathfrak{X}^{1,x}_{t,s}=Y^{0,x}_{t,s}$, let $\mathfrak{Y}^x=(\mathfrak{Y}^x_t)_{t\in[0,T]}\colon[0,T]\times\Omega\to\mathbb{R}^d$, $x\in\mathbb{R}^d$, satisfy for all $x\in\mathbb{R}^d$, $t\in[0,T]$ that $\mathfrak{Y}^x_t=x+\mu(x)t+\sigma(x)W_t$, and let $\tau^x_n\colon\Omega\to[0,T]$, $x\in\mathbb{R}^d$, $n\in\mathbb{N}$, satisfy for all $n\in\mathbb{N}$, $x\in\mathbb{R}^d$ that $\tau^x_n=\inf(\{T\}\cup\{t\in[0,T]\colon[\sup_{s\in[0,t]}\varphi(\mathfrak{Y}^x_s)]+\int_0^t\sum_{i=1}^m|(\varphi'(\mathfrak{Y}^x_s))(\sigma_i(x))|^2\,ds\geq n\})$. Observe that (115) establishes items (i) and (ii) (cf., e.g., Rogers & Williams [54, Theorem 13.1 and Lemma 13.6]). Next note that (115) and (112) imply that for all $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}\max\{\|\mu(x)\|,|||\sigma(x)|||\}&\leq\max\{\|\mu(x)-\mu(0)\|+\|\mu(0)\|,|||\sigma(x)-\sigma(0)|||+|||\sigma(0)|||\}\\&\leq\max\{c\|x\|+\|\mu(0)\|,c\|x\|+|||\sigma(0)|||\}\leq c(\varphi(x))^{\frac{1}{p}}.\end{aligned}\tag{121}$$

This, (112), and the fact that $\forall\,a,b\in[0,\infty),\lambda\in(0,1)\colon a^\lambda b^{1-\lambda}\leq\lambda a+(1-\lambda)b$ imply that for all $x,y\in\mathbb{R}^d$ it holds that

$$\begin{aligned}&\big|(\varphi'(y))(\mu(x))\big|+\frac{1}{2}\left|\sum_{k=1}^m(\varphi''(y))(\sigma_k(x),\sigma_k(x))\right|\\&\leq c(\varphi(y))^{1-\frac{1}{p}}\|\mu(x)\|+\tfrac{c}{2}(\varphi(y))^{1-\frac{2}{p}}\sum_{k=1}^m\|\sigma_k(x)\|^2=c(\varphi(y))^{1-\frac{1}{p}}\|\mu(x)\|+\tfrac{c}{2}(\varphi(y))^{1-\frac{2}{p}}|||\sigma(x)|||^2\\&\leq c(\varphi(y))^{1-\frac{1}{p}}c(\varphi(x))^{\frac{1}{p}}+\tfrac{c}{2}(\varphi(y))^{1-\frac{2}{p}}c^2(\varphi(x))^{\frac{2}{p}}\\&\leq c^2\left[\left(1-\tfrac{1}{p}\right)\varphi(y)+\tfrac{1}{p}\varphi(x)\right]+\tfrac{c^3}{2}\left[\left(1-\tfrac{2}{p}\right)\varphi(y)+\tfrac{2}{p}\varphi(x)\right]\\&\leq\left[c^3\left(1-\tfrac{1}{p}\right)+\tfrac{c^3}{2}\left(1-\tfrac{2}{p}\right)\right]\varphi(y)+\left[\tfrac{c^3}{p}+\tfrac{2c^3}{2p}\right]\varphi(x)=\left(\tfrac{3c^3}{2}-\tfrac{2c^3}{p}\right)\varphi(y)+\tfrac{2c^3}{p}\varphi(x).\end{aligned}\tag{122}$$

Combining this and, e.g., Cox et al. [15, Lemma 2.2] (applied for every $t\in[0,T)$, $s\in[t,T]$, $x\in\mathbb{R}^d$, $\theta\in\Theta$ with $T\curvearrowleft T-t$, $O\curvearrowleft\mathbb{R}^d$, $V\curvearrowleft([0,T-t]\times\mathbb{R}^d\ni(s,x)\mapsto\varphi(x)\in[0,\infty))$, $\alpha\curvearrowleft([0,T-t]\ni s\mapsto 2c^3\in[0,\infty))$, $\tau\curvearrowleft s-t$, $X\curvearrowleft(X^{\theta,x}_{t,t+r})_{r\in[0,T-t]}$ in the notation of Cox et al. [15, Lemma 2.2]) demonstrates that for all $\theta\in\Theta$, $x\in\mathbb{R}^d$, $t\in[0,T]$, $s\in[t,T]$ it holds that

$$\mathbb{E}\big[\varphi(X^{\theta,x}_{t,s})\big]\leq e^{2c^3(s-t)}\varphi(x).\tag{123}$$

Itô's formula, (122), and the fact that $\varphi\geq 1$ imply that for all $x\in\mathbb{R}^d$, $t\in[0,T]$ it holds that

$$\begin{aligned}&\mathbb{E}[\varphi(\mathfrak{Y}^x_{\min\{\tau^x_n,t\}})]\\&=\varphi(x)+\mathbb{E}\left[\int_0^{\min\{\tau^x_n,t\}}(\varphi'(\mathfrak{Y}^x_s))(\mu(x))+\frac{1}{2}\sum_{k=1}^m(\varphi''(\mathfrak{Y}^x_s))(\sigma_k(x),\sigma_k(x))\,ds\right]\\&\leq\varphi(x)+\mathbb{E}\left[\int_0^{\min\{\tau^x_n,t\}}\left(\tfrac{3c^3}{2}-\tfrac{2c^3}{p}\right)\varphi(\mathfrak{Y}^x_s)+\tfrac{2c^3}{p}\varphi(x)\,ds\right]\\&\leq\varphi(x)\left(1+\tfrac{2c^3t}{p}\right)+\left(\tfrac{3c^3}{2}-\tfrac{2c^3}{p}\right)\mathbb{E}\left[\int_0^t\varphi(\mathfrak{Y}^x_s)\mathbb{1}_{[0,\tau^x_n]}(s)\,ds\right]\\&\leq\varphi(x)\left(1+\tfrac{2c^3t}{p}\right)+\left(\tfrac{3c^3}{2}-\tfrac{2c^3}{p}\right)\int_0^t\mathbb{E}[\varphi(\mathfrak{Y}^x_{\min\{\tau^x_n,s\}})]\,ds.\end{aligned}\tag{124}$$

Gronwall's inequality and the fact that for all $a\in\mathbb{R}$ it holds that $1+a\leq e^a$ therefore assure that for all $x\in\mathbb{R}^d$, $t\in[0,T]$ it holds that

$$\mathbb{E}[\varphi(\mathfrak{Y}^x_{\min\{\tau^x_n,t\}})]\leq\exp\!\left(\left[\tfrac{3c^3}{2}-\tfrac{2c^3}{p}\right]t\right)\left[1+\tfrac{2c^3t}{p}\right]\varphi(x)\leq e^{2c^3t}\varphi(x).\tag{125}$$

Fatou's lemma hence proves that for all $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$\mathbb{E}\big[\varphi(x+\mu(x)t+\sigma(x)W_t)\big] = \mathbb{E}[\varphi(\mathfrak{Y}_t^x)] \leq e^{2c^3t}\varphi(x) \tag{126}$$

(cf., e.g., also Hudde et al. [32, Theorem 2.4]). The tower property for conditional expectations, the fact that for all $t \in [0,T]$, $s \in [t,T]$, $\theta \in \Theta$ it holds that $W_s^\theta - W_t^\theta$ and $\mathbb{F}_t$ are independent, and the fact that for all $t \in [0,T]$, $s \in [t,T]$, $\theta \in \Theta$, $B \in \mathcal{B}(\mathbb{R}^d)$ it holds that $\mathbb{P}((W_s^\theta - W_t^\theta) \in B) = \mathbb{P}(W_{s-t}^\theta \in B)$ hence prove that for all $\theta \in \Theta$, $x \in \mathbb{R}^d$, $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[\varphi(Y_{t,s}^{\theta,x})\big] \\
&= \mathbb{E}\Bigg[\mathbb{E}\Big[\varphi\big(Y_{t,\max\{t,\llcorner s\lrcorner\}}^{\theta,x} + \mu(Y_{t,\max\{t,\llcorner s\lrcorner\}}^{\theta,x})(s-\max\{t,\llcorner s\lrcorner\}) \\
&\quad + \sigma(Y_{t,\max\{t,\llcorner s\lrcorner\}}^{\theta,x})(W_s^\theta - W_{\max\{t,\llcorner s\lrcorner\}}^\theta)\big)\Big|\mathbb{F}_{\llcorner s\lrcorner}\Big]\Bigg] \\
&= \mathbb{E}\Bigg[\mathbb{E}\Big[\varphi\big(z+\mu(z)(s-\max\{t,\llcorner s\lrcorner\}) + \sigma(z)(W_{s-\max\{t,\llcorner s\lrcorner\}}^\theta)\big)\Big]\Big|_{z=Y_{t,\max\{t,\llcorner s\lrcorner\}}^{\theta,x}}\Bigg] \\
&\leq e^{2c^3(s-\max\{t,\llcorner s\lrcorner\})}\mathbb{E}\Big[\varphi\big(Y_{t,\max\{t,\llcorner s\lrcorner\}}^{\theta,x}\big)\Big].
\end{aligned} \tag{127}$$

Induction and (116) hence show that for all $\theta \in \Theta$, $x \in \mathbb{R}^d$, $t \in [0,T]$, $s \in [t,T]$ it holds that $\mathbb{E}\big[\varphi(Y_{t,s}^{\theta,x})\big] \leq e^{2c^3(s-t)}\varphi(x)$. Jensen's inequality and (123) therefore prove that for all $q \in [0,p]$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&\max\big\{\mathbb{E}\big[(\varphi(Y_{t,s}^{\theta,x}))^{\frac{q}{p}}\big], \mathbb{E}\big[(\varphi(X_{t,s}^{\theta,x}))^{\frac{q}{p}}\big]\big\} \\
&\leq \max\bigg\{\Big(\mathbb{E}\big[\varphi(Y_{t,s}^{\theta,x})\big]\Big)^{\frac{q}{p}}, \Big(\mathbb{E}\big[\varphi(X_{t,s}^{\theta,x})\big]\Big)^{\frac{q}{p}}\bigg\} \leq e^{2qc^3(s-t)/p}(\varphi(x))^{\frac{q}{p}}.
\end{aligned} \tag{128}$$

Moreover, observe that the fact that μ is continuous, the fact that σ is continuous, the fact that for all $\theta \in \Theta$, $\omega \in \Omega$ it holds that $[0,T] \ni t \mapsto W_t^\theta(\omega) \in \mathbb{R}^d$ is continuous, and Fubini's theorem imply that for all $\theta \in \Theta$ and all measurable $\eta\colon [0,T] \times \mathbb{R}^d \to [0,\infty)$ it holds that

$$\Delta \times \mathbb{R}^d \ni (t,s,x) \mapsto \mathbb{E}\big[\eta\big(s, Y_{t,s}^{\theta,x}\big)\big] \in [0,\infty] \tag{129}$$

is measurable. Furthermore, note that (112), (115), (122), and, e.g., Beck et al. [4, Lemma 3.7] (applied with $\mathcal{O} \curvearrowleft \mathbb{R}^d$, $V \curvearrowleft ([0,T] \times \mathbb{R}^d \ni (t,x) \mapsto e^{-2c^3t/p}\varphi(x) \in (0,\infty))$ in the notation of Beck et al. [4, Lemma 3.7]) imply that $\Delta \times \mathbb{R}^d \times \mathbb{R}^d \ni (t,s,x,y) \mapsto \big(s, X_{t,s}^{\theta,x}, X_{t,s}^{\theta,y}\big) \in \mathcal{L}^0(\Omega; \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d)$ is continuous. This and the dominated convergence theorem prove that for all $\theta \in \Theta$ and all bounded and continuous $\eta\colon [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ it holds that $\Delta \times \mathbb{R}^d \times \mathbb{R}^d \ni (t,s,x,y) \mapsto \mathbb{E}\big[\eta\big(s, X_{t,s}^{\theta,x}, X_{t,s}^{\theta,y}\big)\big] \in [0,\infty]$ is continuous. Hence, we obtain that for all $\theta \in \Theta$ and all bounded and continuous $\eta\colon [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ it holds that $\Delta \times \mathbb{R}^d \times \mathbb{R}^d \ni (t,s,x,y) \mapsto \mathbb{E}\big[\eta\big(s, X_{t,s}^{\theta,x}, X_{t,s}^{\theta,y}\big)\big] \in [0,\infty]$ is measurable. This implies that for all $\theta \in \Theta$ and all measurable $\eta\colon [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ it holds that

$$\Delta \times \mathbb{R}^d \times \mathbb{R}^d \ni (t,s,x,y) \mapsto \mathbb{E}\big[\eta\big(s, X_{t,s}^{\theta,x}, X_{t,s}^{\theta,y}\big)\big] \in [0,\infty] \tag{130}$$

is measurable. Combining (129), (128), (113), (114), and Proposition 2.2 (applied for every $k \in \{0,1\}$ with $L \curvearrowleft c$, $\mathcal{O} \curvearrowleft \mathbb{R}^d$, $(X_{t,s}^x)_{(t,s,x)\in\Delta\times\mathbb{R}^d} \curvearrowleft (\mathfrak{X}_{t,s}^{k,x})_{(t,s,x)\in\Delta\times\mathbb{R}^d}$, $V \curvearrowleft ([0,T] \times \mathbb{R}^d \ni (s,x) \mapsto e^{2c^3\beta(T-s)/p}(\varphi(x))^{\beta/p} \in (0,\infty))$ in the notation of Proposition 2.2) hence establishes that

a) there exist unique measurable $u_k\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $k \in \{0,1\}$, which satisfy for all $k \in \{0,1\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\sup_{s\in[0,T]}\sup_{x\in\mathbb{R}^d}\big[|u_k(s,x)|(\varphi(x))^{-\beta/p}\big] + \mathbb{E}\big[\big|g\big(\mathfrak{X}_{t,T}^{k,x}\big)\big| + \int_t^T \big|f\big(s, \mathfrak{X}_{t,s}^{k,x}, u_k(s, \mathfrak{X}_{t,s}^{k,x})\big)\big|\,ds\big] < \infty$ and

$$u_k(t,x) = \mathbb{E}\bigg[g\big(\mathfrak{X}_{t,T}^{k,x}\big) + \int_t^T f\big(s, \mathfrak{X}_{t,s}^{k,x}, u_k(s, \mathfrak{X}_{t,s}^{k,x})\big)\,ds\bigg] \tag{131}$$

and

b) it holds for all $k \in \{0,1\}$ that

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left[\frac{|\mathfrak{u}_k(t,x)|}{e^{2c^3\beta(T-t)/p}(\varphi(x))^{\beta/p}}\right] \le \sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left[\left[\frac{|g(x)|}{(\varphi(x))^{\beta/p}}+\frac{|Tf(t,x,0)|}{(\varphi(x))^{\beta/p}}\right]e^{cT}\right] \le 2be^{cT}. \tag{132}$$

This proves item (iii). Moreover, note that Lemma 3.2 establishes item (iv). Next observe that (121) and (128) demonstrate that for all $\theta\in\Theta$, $t\in[0,T]$, $r\in[t,T]$, $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}&\max\left\{\mathbb{E}\Big[\big\|\mu(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})\big\|^2\Big],\mathbb{E}\Big[\big\|\!\big|\sigma(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})\big|\!\big\|^2\Big]\right\}\\&\le c^2\mathbb{E}\Big[\big(\varphi(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})\big)^{2/p}\Big]\le c^2e^{4c^3(r-t)/p}(\varphi(x))^{2/p}.\end{aligned} \tag{133}$$

Furthermore, note that (116) demonstrates that for all $t\in[0,T]$, $r\in[t,T]$, $x\in\mathbb{R}^d$, $\theta\in\Theta$ it holds that $\sigma(\{Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}}\})\subseteq\mathbb{F}_r$. Combining this and (133) with the fact that for all $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that $\mathbb{E}\big[\|\sigma(x)W_t\|^2\big]=|\!|\!|\sigma(x)|\!|\!|^2t$ shows that for all $\theta\in\Theta$, $t\in[0,T]$, $r\in[t,T]$, $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}&\mathbb{E}\Big[\big\|\sigma(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})(W^\theta_r-W^\theta_{\max\{t,\llcorner r\lrcorner\}})\big\|^2\Big]\\&=\mathbb{E}\left[\mathbb{E}\Big[\big\|\sigma(y)(W^\theta_r-W^\theta_{\max\{t,\llcorner r\lrcorner\}})\big\|^2\Big]\Big|_{y=Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}}}\right]\\&=\mathbb{E}\Big[|\!|\!|\sigma(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})|\!|\!|^2(r-\max\{t,\llcorner r\lrcorner\})\Big]\\&\le\mathbb{E}\Big[|\!|\!|\sigma(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})|\!|\!|^2\delta\Big]\le c^2e^{4c^3(r-t)/p}(\varphi(x))^{2/p}\delta.\end{aligned} \tag{134}$$

This, (116), the triangle inequality, and (133) imply that for all $\theta\in\Theta$, $t\in[0,T]$, $r\in[t,T]$, $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}&\left(\mathbb{E}\Big[\big\|Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}}-Y^{\theta,x}_{t,r}\big\|^2\Big]\right)^{1/2}\\&\le\left(\mathbb{E}\Big[\big\|\mu(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})\big\|^2\Big]\right)^{1/2}(r-\max\{t,\llcorner r\lrcorner\})\\&\quad+\left(\mathbb{E}\Big[\big\|\sigma(Y^{\theta,x}_{t,\max\{t,\llcorner r\lrcorner\}})(W^\theta_r-W^\theta_{\max\{t,\llcorner r\lrcorner\}})\big\|^2\Big]\right)^{1/2}\\&\le ce^{2c^3(r-t)/p}(\varphi(x))^{1/p}\delta^{1/2}|r-t|^{1/2}+ce^{2c^3(r-t)/p}(\varphi(x))^{1/p}\delta^{1/2}\\&=c\big[|r-t|^{1/2}+1\big]e^{2c^3(r-t)/p}(\varphi(x))^{1/p}\delta^{1/2}.\end{aligned} \tag{135}$$

Next note that (115) and the fact that $c\ge1$ assure that for all $z,y\in\mathbb{R}^d$ with $z\ne y$ it holds that

$$\frac{\langle z-y,\mu(z)-\mu(y)\rangle+\frac{1}{2}|\!|\!|\sigma(z)-\sigma(y)|\!|\!|^2}{\|z-y\|^2}+\frac{(\frac{2}{2}-1)\|(\sigma(z)-\sigma(y))^\intercal(z-y)\|^2}{\|z-y\|^4}\le 2c^2. \tag{136}$$

This, [35, Theorem 1.2] (applied for every $\theta\in\Theta$, $t\in[0,T)$, $s\in(t,T]$, $x\in\mathbb{R}^d$ with $H\curvearrowleft\mathbb{R}^d$, $U\curvearrowleft\mathbb{R}^m$, $D\curvearrowleft\mathbb{R}^d$, $T\curvearrowleft(s-t)$, $(\mathbb{F}_r)_{r\in[0,T]}\curvearrowleft(\mathbb{F}_{r+t})_{r\in[0,s-t]}$, $(W_r)_{r\in[0,T]}\curvearrowleft(W^\theta_{t+r}-W^\theta_t)_{r\in[0,s-t]}$, $(X_r)_{r\in[0,T]}\curvearrowleft(X^{\theta,x}_{t,t+r})_{r\in[0,s-t]}$, $(Y_r)_{r\in[0,T]}\curvearrowleft(Y^{\theta,x}_{t,t+r})_{r\in[0,s-t]}$, $(a_r)_{r\in[0,T]}\curvearrowleft(\mu(Y^{\theta,x}_{t,\max\{t,\llcorner t+r\lrcorner\}}))_{r\in[0,s-t]}$, $(b_r)_{r\in[0,T]}\curvearrowleft(\sigma(Y^{\theta,x}_{t,\max\{t,\llcorner t+r\lrcorner\}}))_{r\in[0,s-t]}$, $\varepsilon\curvearrowleft1$, $p\curvearrowleft2$, $\tau\curvearrowleft(\Omega\ni\omega\mapsto s-t\in[0,s-t])$, $\alpha\curvearrowleft1$, $\beta\curvearrowleft1$, $r\curvearrowleft2$, $q\curvearrowleft\infty$ in the notation of [35, Theorem 1.2]), (115), (135), the fact that for all $t\in[0,\infty)$ it holds that $\sqrt{t}(\sqrt{t}+1)\le e^t$, the fact that $1\le c$, and the fact that $p \geq 2$ imply that for all $\theta \in \Theta$, $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{split} &\Big(\mathbb{E}\Big[\big\|X^{\theta,x}_{t,s} - Y^{\theta,x}_{t,s}\big\|^2\Big]\Big)^{1/2} \\ &\leq \sup_{\substack{z,y\in\mathbb{R}^d,\\ z\neq y}} \exp\left(\int_t^s \left[\frac{\langle z-y, \mu(z)-\mu(y)\rangle + \frac{(2-1)(1+1)}{2}\|\sigma(z)-\sigma(y)\|^2}{\|z-y\|^2} + \tfrac{1-\frac{1}{2}}{1} + \tfrac{\frac{1}{2}-\frac{1}{2}}{1}\right]^+ dr\right) \\ &\quad \cdot \left[\left(\int_t^s \mathbb{E}\Big[\big\|\mu\big(Y^{\theta,x}_{t,\max\{t,\llcorner r \lrcorner\}}\big) - \mu\big(Y^{\theta,x}_{t,r}\big)\big\|^2\Big]\,dr\right)^{1/2} \right. \\ &\quad \left. + \sqrt{\tfrac{(2-1)(1+1)}{1}}\left(\int_t^s \mathbb{E}\Big[\big\|\sigma\big(Y^{\theta,x}_{t,\max\{t,\llcorner r \lrcorner\}}\big) - \sigma\big(Y^{\theta,x}_{t,r}\big)\big\|^2\Big]\,dr\right)^{1/2}\right] \\ &\leq e^{3c^2(s-t)} 3c \left(|s-t| \sup_{r\in[t,s]} \mathbb{E}\Big[\big\|Y^{\theta,x}_{t,\max\{t,\llcorner r \lrcorner\}} - Y^{\theta,x}_{t,r}\big\|^2\Big]\right)^{1/2} \\ &\leq e^{3c^2(s-t)} 3c|s-t|^{1/2} c\big[|s-t|^{1/2}+1\big] e^{2c^3(s-t)/p}(\varphi(x))^{1/p}\delta^{1/2} \\ &\leq 3c^2 e^{4c^2T} e^{2c^3(s-t)/p}(\varphi(x))^{1/p}\delta^{1/2}. \end{split} \tag{137}$$

Next observe that item (i), (136), and, e.g., Cox et al. [15, Corollary 2.26] (applied for every $t \in [0,T)$, $s \in (t,T]$ with $T \curvearrowleft s-t$, $O \curvearrowleft \mathbb{R}^d$, $(\mathcal{F}_r)_{r\in[0,T]} \curvearrowleft (\mathbb{F}_{t,t+r})_{r\in[0,s-t]}$, $(W_r)_{r\in[0,T]} \curvearrowleft (W^0_{t,t+r} - W^0_t)_{r\in[0,s-t]}$, $\alpha_0 \curvearrowleft 0$, $\alpha_1 \curvearrowleft 0$, $\beta_0 \curvearrowleft 0$, $\beta_1 \curvearrowleft 0$, $c \curvearrowleft 2c^2$, $r \curvearrowleft 2$, $p \curvearrowleft 2$, $q_0 \curvearrowleft \infty$, $q_1 \curvearrowleft \infty$, $U_0 \curvearrowleft (\mathbb{R}^d \ni x \mapsto 0 \in \mathbb{R})$, $U_1 \curvearrowleft (\mathbb{R}^d \ni x \mapsto 0 \in [0,\infty))$, $\overline{U} \curvearrowleft (\mathbb{R}^d \ni x \mapsto 0 \in \mathbb{R})$, $(X^x_r)_{r\in[0,T],x\in\mathbb{R}^d} \curvearrowleft (X^{0,x}_{t,t+r})_{r\in[0,s-t],x\in\mathbb{R}^d}$ in the notation of Cox et al. [15, Corollary 2.26]) demonstrate that for all $t \in [0,T)$, $s \in (t,T]$, $x,y \in \mathbb{R}^d$ it holds that $\big(\mathbb{E}\big[\|X^{0,x}_{t,s} - X^{0,y}_{t,s}\|^2\big]\big)^{1/2} \leq e^{2c^2(s-t)}\|x-y\|$. This and (137) imply that for all $t \in [0,T]$, $s \in [t,T]$, $r \in [s,T]$, $x,y \in \mathbb{R}^d$ it holds that

$$\begin{split} &\left(\mathbb{E}\left[\mathbb{E}\Big[\big\|X^{0,\mathfrak{x}}_{s,r} - X^{0,\mathfrak{y}}_{s,r}\big\|^2\Big]\Big|_{(\mathfrak{x},\mathfrak{y})=(X^{0,x}_{t,s},Y^{0,x}_{t,s})}\right]\right)^{1/2} \leq \left(\mathbb{E}\left[\Big[e^{2c^2(r-s)}\,\big\|X^{0,x}_{t,s} - Y^{0,x}_{t,s}\big\|\Big]^2\right]\right)^{1/2} \\ &\leq e^{2c^2(r-s)} 3c^2 e^{4c^2T} e^{2c^3(s-t)/p}(\varphi(x))^{1/p}\delta^{1/2} \leq 3c^2 e^{4c^2T}\delta^{1/2}\big[e^{4c^3(T-t)/p}(\varphi(x))^{2/p}\big]^{1/2}. \end{split} \tag{138}$$

Furthermore, note that item (i) and Tonelli's theorem ensure that for all $t \in [0,T]$, $s \in [t,T]$, $r \in [s,T]$, $x,y \in \mathbb{R}^d$ and all measurable $h\colon \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ it holds that $\mathbb{R}^d \times \mathbb{R}^d \ni (y_1,y_2) \mapsto \mathbb{E}\big[h\big(X^{0,y_1}_{s,r}, X^{0,y_2}_{s,r}\big)\big] \in [0,\infty]$ is measurable. Moreover, observe that item (i) assures that for all $t \in [0,T]$, $s \in [t,T]$, $r \in [s,T]$, $x,y \in \mathbb{R}^d$ it holds that $X^{0,x}_{t,s}$ and $X^{0,y}_{s,r}$ are independent. This and, e.g., the disintegration-type result in [38, Lemma 2.2] show that for all $t \in [0,T]$, $s \in [t,T]$, $r \in [s,T]$, $x,y \in \mathbb{R}^d$ and all measurable $h\colon \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ it holds that $\mathbb{E}\big[\mathbb{E}\big[h\big(X^{0,\tilde{x}}_{s,r}, X^{0,\tilde{y}}_{s,r}\big)\big]|_{\tilde{x}=X^{0,x}_{t,s},\tilde{y}=X^{0,y}_{t,s}}\big] = \mathbb{E}\big[h\big(X^{0,x}_{t,r}, X^{0,y}_{t,r}\big)\big]$. Combining item (i), (116), (128), (130), (114), (138), (131), (132), Lemma 2.3 (applied with $L \curvearrowleft c$, $\rho \curvearrowleft 2c^3$, $\eta \curvearrowleft 1$, $\delta \curvearrowleft 3c^2e^{4c^2T}\delta^{1/2}$, $p \curvearrowleft p/\beta$, $q \curvearrowleft 2$, $(X^{x,1}_{t,s})_{t\in[0,T],s\in[t,T],x\in\mathbb{R}^d} \curvearrowleft (X^{0,x}_{t,s})_{t\in[0,T],s\in[t,T],x\in\mathbb{R}^d}$, $(X^{x,2}_{t,s})_{t\in[0,T],s\in[t,T],x\in\mathbb{R}^d} \curvearrowleft (Y^{0,x}_{t,s})_{t\in[0,T],s\in[t,T],x\in\mathbb{R}^d}$, $V \curvearrowleft b^{p/\beta}\varphi$, $\psi \curvearrowleft \big([0,T] \times \mathbb{R}^d \ni (t,x) \mapsto e^{4c^3(T-t)/p}(\varphi(x))^{2/p} \in (0,\infty)\big)$, $u_1 \curvearrowleft \mathcal{u}_0$, $u_2 \curvearrowleft \mathcal{u}_1$ in the notation of Lemma 2.3), the fact that $1 + cT \leq e^{cT}$, the fact that $c \geq 1$, the fact that $\varphi \geq 1$, the fact that $p \geq 2$, and the fact that $p \geq 2\beta$ hence implies that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{split} &|\mathcal{u}_0(t,x) - \mathcal{u}_1(t,x)| \\ &\leq 4(1+cT)T^{-1/2}e^{cT+(2c^3\beta/p+c)T}\big(b^{p/\beta}\varphi(x)\big)^{\beta/p}\big[e^{4c^3(T-t)/p}(\varphi(x))^{2/p}\big]^{1/2} 3c^2e^{4c^2T}\delta^{1/2} \\ &\leq 4e^{cT}T^{-1/2}e^{cT+c^3T+cT}b(\varphi(x))^{\frac{\beta}{p}}e^{c^3T}(\varphi(x))^{\frac{1}{p}}3c^2e^{4c^2T}\delta^{1/2} \\ &\leq 12bc^2T^{-1/2}e^{9c^3T}(\varphi(x))^{\frac{\beta+1}{p}}\delta^{1/2}. \end{split} \tag{139}$$

Next observe that Corollary 3.12 (applied for every $t \in [0,T)$, $n \in \mathbb{N}$ with $L \curvearrowleft c$, $\rho \curvearrowleft 4\beta c^3/p$, $Y^\theta \curvearrowleft Y^\theta$, $U^\theta \curvearrowleft U^\theta$, $u \curvearrowleft \mathcal{u}_1$, $\varphi \curvearrowleft \varphi^{2\beta/p}$, $N \curvearrowleft n$, $t_0 \curvearrowleft t$ in the notation of Corollary 3.12), (114), (131), (132), (128), (113), and the fact that $p \geq 2\beta \geq 2$ assure that for all $t \in [0,T]$, $n \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d}\left[\frac{\mathbb{E}[|U_n^0(t,x)-\mathfrak{u}_1(t,x)|^2]}{(\varphi(x))^{2\beta/p}}\right]^{1/2} \\
&\leq e^{M/2}M^{-n/2}(1+2Tc)^{n-1}e^{2c^3\beta T/p} \\
&\quad\cdot\left[2\sup_{s\in[0,T]}\sup_{x\in\mathbb{R}^d}\left[\frac{\max\{|T(F(0))(s,x)|,|g(x)|\}}{(\varphi(x))^{\beta/p}}\right]+2Tc\sup_{s\in[0,T]}\sup_{x\in\mathbb{R}^d}\left[\frac{|\mathfrak{u}_1(s,x)|}{(\varphi(x))^{\beta/p}}\right]\right] \\
&\leq e^{M/2}M^{-n/2}(1+2Tc)^{n-1}e^{2c^3\beta T/p}\left[2b+4Tcbe^{cT+2c^3\beta T/p}\right] \\
&\leq e^{M/2}M^{-n/2}(1+2Tc)^{n-1}e^{2c^3\beta T/p}2be^{cT+2c^3\beta T/p}(1+2Tc)\leq 2be^{M/2}M^{-n/2}e^{2ncT}e^{3c^3T}.
\end{aligned}\tag{140}$$

The triangle inequality, (139), the fact that $c \geq 1$, the fact that $\varphi \geq 1$, and the fact that $p \geq 2$ hence show that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\left(\mathbb{E}\left[|U_n^0(t,x)-\mathfrak{u}_0(t,x)|^2\right]\right)^{1/2}\leq\left(\mathbb{E}\left[|U_n^0(t,x)-\mathfrak{u}_1(t,x)|^2\right]\right)^{1/2}+|\mathfrak{u}_1(t,x)-\mathfrak{u}_0(t,x)| \\
&\leq 2be^{M/2}M^{-n/2}e^{2ncT}e^{3c^3T}(\varphi(x))^{\frac{\beta}{p}}+12bc^2T^{-1/2}e^{9c^3T}(\varphi(x))^{\frac{\beta+1}{p}}\delta^{1/2} \\
&\leq\left(e^{M/2}e^{2ncT}M^{-n/2}+\delta^{1/2}T^{-1/2}\right)12bc^2e^{9c^3T}(\varphi(x))^{\frac{\beta+1}{p}}.
\end{aligned}\tag{141}$$

This and (132) establish item (v). The proof of Proposition 4.1 is thus complete. □

4.2 Complexity analysis for MLP approximations in fixed space dimensions

Theorem 4.2. *Let $d,m \in \mathbb{N}$, $T \in (0,\infty)$, $\mathfrak{f},\mathfrak{g},\mathfrak{m} \in [0,\infty)$, $\beta,b,c \in [1,\infty)$, $p \in [2\beta,\infty)$, $\varphi \in C^2(\mathbb{R}^d,[1,\infty))$, $g \in C(\mathbb{R}^d,\mathbb{R})$, $\mu \in C(\mathbb{R}^d,\mathbb{R}^d)$, $\sigma = (\sigma_1,\sigma_2,\ldots,\sigma_m) \in C(\mathbb{R}^d,\mathbb{R}^{d\times m})$, let $\|\cdot\|\colon \mathbb{R}^d \to [0,\infty)$ be the standard norm on $\mathbb{R}^d$, let $f\colon [0,T]\times\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$ be measurable, let $F\colon \mathbb{R}^{[0,T]\times\mathbb{R}^d}\to\mathbb{R}^{[0,T]\times\mathbb{R}^d}$ satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $v\in\mathbb{R}^{[0,T]\times\mathbb{R}^d}$ that $(F(v))(t,x) = f(t,x,v(t,x))$, assume for all $x,y\in\mathbb{R}^d$, $z\in\mathbb{R}^d\backslash\{0\}$, $t\in[0,T]$, $v,w\in\mathbb{R}$ that*

$$\max\left\{\tfrac{|(\varphi'(x))(z)|}{(\varphi(x))^{(p-1)/p}\|z\|},\tfrac{|(\varphi''(x))(z,z)|}{(\varphi(x))^{(p-2)/p}\|z\|^2},\tfrac{c\|x\|+\|\mu(0)\|}{(\varphi(x))^{1/p}},\tfrac{c\|x\|+[\sum_{i=1}^m\|\sigma_i(0)\|^2]^{1/2}}{(\varphi(x))^{1/p}}\right\}\leq c, \tag{142}$$

$$\max\left\{|Tf(t,x,0)|,|g(x)|\right\}\leq b(\varphi(x))^{\beta/p}, \tag{143}$$

$$\max\left\{|g(x)-g(y)|,T|f(t,x,v)-f(t,y,w)|\right\}\leq cT|v-w|+\tfrac{(\varphi(x)+\varphi(y))^{\beta/p}\|x-y\|}{T^{1/2}b^{-1}}, \tag{144}$$

$$\max\left\{\|\mu(x)-\mu(y)\|^2,\textstyle\sum_{i=1}^m\|\sigma_i(x)-\sigma_i(y)\|^2\right\}\leq c^2\|x-y\|^2, \tag{145}$$

let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $\Theta = \bigcup_{n\in\mathbb{N}}\mathbb{Z}^n$, let $\mathfrak{r}^\theta\colon\Omega\to[0,1]$, $\theta\in\Theta$, be i.i.d. random variables, assume for all $t\in(0,1)$ that $\mathbb{P}(\mathfrak{r}^0\leq t)=t$, let $\mathcal{R}^\theta\colon[0,T]\times\Omega\to[0,T]$, $\theta\in\Theta$, satisfy for all $\theta\in\Theta$, $t\in[0,T]$ that $\mathcal{R}^\theta_t = t+(T-t)\mathfrak{r}^\theta$, let $W^\theta\colon[0,T]\times\Omega\to\mathbb{R}^m$, $\theta\in\Theta$, be i.i.d. standard $(\mathbb{F}_t)_{t\in[0,T]}$-Brownian motions, assume that $(\mathfrak{r}^\theta)_{\theta\in\Theta}$ and $(W^\theta)_{\theta\in\Theta}$ are independent, for every $N\in\mathbb{N}$, $\theta\in\Theta$, $x\in\mathbb{R}^d$, $t\in[0,T]$ let $Y^{N,\theta,x}_t = (Y^{N,\theta,x}_{t,s})_{s\in[t,T]}\colon[t,T]\times\Omega\to\mathbb{R}^d$ satisfy for all $n\in\{0,1,\ldots,N\}$, $s\in[\frac{nT}{N},\frac{(n+1)T}{N}]\cap[t,T]$ that $Y^{N,\theta,x}_{t,t}=x$ and

$$\begin{aligned}
&Y^{N,\theta,x}_{t,s}-Y^{N,\theta,x}_{t,\max\{t,nT/N\}} \\
&=\mu\big(Y^{N,\theta,x}_{t,\max\{t,nT/N\}}\big)\big(s-\max\{t,\tfrac{nT}{N}\}\big)+\sigma\big(Y^{N,\theta,x}_{t,\max\{t,nT/N\}}\big)\big(W^\theta_s-W^\theta_{\max\{t,nT/N\}}\big),
\end{aligned}\tag{146}$$

let $U^\theta_{n,M}\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $n,M\in\mathbb{Z}$, $\theta\in\Theta$, *satisfy for all* $\theta\in\Theta$, $n\in\mathbb{N}_0$, $M\in\mathbb{N}$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *that*

$$\begin{aligned} U^\theta_{n,M}(t,x) &= \frac{\mathbb{1}_{\mathbb{N}}(n)}{M^n}\sum_{i=1}^{M^n} g\big(Y^{M^M,(\theta,0,-i),x}_{t,T}\big) \\ &+ \sum_{\ell=0}^{n-1}\frac{(T-t)}{M^{n-\ell}}\left[\sum_{i=1}^{M^{n-\ell}}\big(F\big(U^{(\theta,\ell,i)}_{\ell,M}\big)-\mathbb{1}_{\mathbb{N}}(\ell)F\big(U^{(\theta,-\ell,i)}_{\ell-1,M}\big)\big)\big(\mathcal{R}^{(\theta,\ell,i)}_t, Y^{M^M,(\theta,\ell,i),x}_{t,\mathcal{R}^{(\theta,\ell,i)}_t}\big)\right], \end{aligned} \tag{147}$$

and let $\mathfrak{C}_{n,M}\in\mathbb{R}$, $n,M\in\mathbb{Z}$, *satisfy for all* $n\in\mathbb{Z}$, $M\in\mathbb{N}$ *that*

$$\mathfrak{C}_{n,M}\le M^n(M^M\mathfrak{m}+\mathfrak{g})\mathbb{1}_{\mathbb{N}}(n)+\sum_{\ell=0}^{n-1}\big[M^{n-\ell}(M^M\mathfrak{m}+\mathfrak{f}+\mathfrak{C}_{\ell,M}+\mathfrak{C}_{\ell-1,M})\big]. \tag{148}$$

Then

(i) *for every* $t\in[0,T]$, $x\in\mathbb{R}^d$, $\theta\in\Theta$ *there exists a unique* $(\mathbb{F}_s)_{s\in[t,T]}$*-adapted stochastic process* $X^{\theta,x}_t=(X^{\theta,x}_{t,s})_{s\in[t,T]}\colon[t,T]\times\Omega\to\mathbb{R}^d$ *with continuous sample paths which satisfies that for all* $s\in[t,T]$ *it holds* $\mathbb{P}$*-a.s. that*

$$X^{\theta,x}_{t,s}=x+\int_t^s\mu(X^{\theta,x}_{t,r})\,dr+\int_t^s\sigma(X^{\theta,x}_{t,r})\,dW^\theta_r, \tag{149}$$

(ii) *there exists a unique measurable* $u\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$ *which satisfies for all* $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $\big(\sup_{s\in[0,T],y\in\mathbb{R}^d}[|u(s,y)|(\varphi(y))^{-\beta/p}]\big)+\int_t^T\mathbb{E}\big[|f(s,X^{0,x}_{t,s},u(s,X^{0,x}_{t,s}))|\big]\,ds+\mathbb{E}\big[|g(X^{0,x}_{t,T})|\big]<\infty$ *and*

$$u(t,x)=\mathbb{E}\big[g(X^{0,x}_{t,T})\big]+\int_t^T\mathbb{E}\big[f(s,X^{0,x}_{t,s},u(s,X^{0,x}_{t,s}))\big]\,ds, \tag{150}$$

(iii) *it holds for all* $t\in[0,T]$, $x\in\mathbb{R}^d$, $n\in\mathbb{N}_0$, $M\in\mathbb{N}$, $\theta\in\Theta$ *that* $U^\theta_{n,M}(t,x)$ *is measurable,*

(iv) *it holds for all* $t\in[0,T]$, $x\in\mathbb{R}^d$, $n\in\mathbb{N}_0$, $M\in\mathbb{N}$ *that*

$$\big(\mathbb{E}\big[|U^0_{n,M}(t,x)-u(t,x)|^2\big]\big)^{1/2}\le\left[\frac{\exp(2ncT+\frac{M}{2})}{M^{n/2}}+\frac{1}{M^{M/2}}\right]12bc^2|\varphi(x)|^{\frac{\beta+1}{p}}\exp(9c^3T), \tag{151}$$

(v) *it holds for all* $n\in\mathbb{N}$ *that* $\sum_{k=1}^{n+1}\mathfrak{C}_{k,k}\le 12(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})36^n n^{2n}$, *and*

(vi) *there exist* $\mathsf{n}\colon(0,1]\times\mathbb{R}^d\to\mathbb{N}$ *such that for all* $\varepsilon,\gamma\in(0,1]$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *it holds that* $\sup_{n\in[\mathsf{n}(\varepsilon,x),\infty)\cap\mathbb{N}}(\mathbb{E}[|U^0_{n,n}(t,x)-u(t,x)|^2])^{1/2}<\varepsilon$ *and*

$$\begin{aligned}\left[\sum_{n=1}^{\mathsf{n}(\varepsilon,x)}\mathfrak{C}_{n,n}\right]\varepsilon^{\gamma+4}&\le(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})\big[\sup_{n\in\mathbb{N}}\big[n^{-\gamma n/2}\,(5^n\exp(2ncT))^{\gamma+4}\big]\big]\\ &\quad\cdot\big[45bc^2\exp(9c^3T)(\varphi(x))^{(\beta+1)/p}\big]^{\gamma+4}<\infty.\end{aligned} \tag{152}$$

Proof of Theorem 4.2. Throughout this proof let $\mathsf{n}\colon(0,1]\times\mathbb{R}^d\to[1,\infty]$ satisfy for all $\varepsilon\in(0,1]$, $x\in\mathbb{R}^d$ that

$$\mathsf{n}(\varepsilon,x)=\inf\left(\left\{n\in\mathbb{N}\colon\sup_{k\in[n,\infty)\cap\mathbb{N}}\sup_{t\in[0,T]}\mathbb{E}\Big[\big|U^0_{k,k}(t,x)-u(t,x)\big|^2\Big]<\varepsilon^2\right\}\cup\{\infty\}\right). \tag{153}$$

Observe that Proposition 4.1 (applied for every $M \in \mathbb{N}$ with $K \curvearrowleft M^M$, $(\tau_k)_{k\in\{0,1,\dots,K\}} \curvearrowleft (\frac{kT}{M^M})_{k\in\{0,1,\dots,M^M\}}$ in the notation of Proposition 4.1) establishes items (i)–(iv). Next note that the fact that $\lim_{n\to\infty}(e^{n/2}e^{2ncT}n^{-n/2}) = 0$ and item (iv) show that for all $x \in \mathbb{R}^d$, $\varepsilon \in (0,1]$ it holds that

$$\mathsf{n}(\varepsilon, x) \in \mathbb{N}. \tag{154}$$

Moreover, observe that (148) and, e.g., Beck et al. [5, Lemma 3.14] (applied for every $M \in \mathbb{N}$ with $\alpha \curvearrowleft (2M^M\mathfrak{m} + \mathfrak{g} + \mathfrak{f})$, $\beta \curvearrowleft (M^M\mathfrak{m} + \mathfrak{f})$, $(C_n)_{n\in\mathbb{N}_0} \curvearrowleft (\mathfrak{C}_{n,M})_{n\in\mathbb{N}_0}$ in the notation of Beck et al. [5, Lemma 3.14]) demonstrate that for all $n, M \in \mathbb{N}$ it holds that

$$\mathfrak{C}_{n,M} \leq \left[\frac{3M^M\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f}}{2}\right](3M)^n. \tag{155}$$

This implies that for all $n \in \mathbb{N}$, $k \in \{1, 2, \dots, n\}$ it holds that

$$\mathfrak{C}_{k,k} \leq \tfrac{(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})(3k^2)^k}{2} \leq \tfrac{(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})(3(n+1)^2)^{n+1}}{2} \leq \tfrac{(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})(3(2n)^2)^{n+1}}{2} = \tfrac{(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})(12n^2)^{n+1}}{2}. \tag{156}$$

The fact that for all $n \in \mathbb{N}$ it holds that $n^3 \leq 3^n$ hence ensures that for all $n \in \mathbb{N}$ it holds that

$$\sum_{k=1}^{n+1} \mathfrak{C}_{k,k} \leq \tfrac{(3\mathfrak{m}+\mathfrak{g}+2\mathfrak{f})(n+1)(12n^2)^{n+1}}{2} \leq (3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})n(12n^2)^{n+1} \leq 12(3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})36^n n^{2n}. \tag{157}$$

This establishes item (v). Next observe that item (iv) and item (v) prove that for all $\gamma \in (0,1]$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\left[\sum_{k=1}^{n+1} \mathfrak{C}_{k,k}\right] \left(\mathbb{E}\left[\left|U^0_{n,n}(t,x) - u(t,x)\right|^2\right]\right)^{\frac{4+\gamma}{2}} \\
&\leq 12(3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})36^n n^{2n} \left(e^{n/2}e^{2ncT}n^{-n/2} + n^{-n/2}\right)^{\gamma+4} \left[12bc^2e^{9c^3T}(\varphi(x))^{(\beta+1)/p}\right]^{\gamma+4} \\
&= 12(3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})36^n n^{-\gamma n/2} \left(e^{n/2}e^{2ncT} + 1\right)^{\gamma+4} \left[12bc^2e^{9c^3T}(\varphi(x))^{(\beta+1)/p}\right]^{\gamma+4} \\
&\leq 12(3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})n^{-\gamma n/2} \left(36^{n/4}e^{n/2}e^{2ncT}\right)^{\gamma+4} \left[24bc^2e^{9c^3T}(\varphi(x))^{(\beta+1)/p}\right]^{\gamma+4} \\
&\leq (3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})n^{-\gamma n/2} \left(5^n e^{2ncT}\right)^{\gamma+4} \left[45bc^2e^{9c^3T}(\varphi(x))^{(\beta+1)/p}\right]^{\gamma+4}.
\end{aligned} \tag{158}$$

This, (153), and (154) show that for all $\varepsilon, \gamma \in (0,1]$, $t \in [0,T]$, $x \in \mathbb{R}^d$ with $\mathsf{n}(\varepsilon,x) \geq 2$ it holds that

$$\begin{aligned}
&\left[\sum_{k=1}^{\mathsf{n}(\varepsilon,x)} \mathfrak{C}_{k,k}\right] \varepsilon^{4+\gamma} \leq \left[\sum_{k=1}^{\mathsf{n}(\varepsilon,x)} \mathfrak{C}_{k,k}\right] \left(\mathbb{E}\left[\left|U^0_{\mathsf{n}(\varepsilon,x)-1,\mathsf{n}(\varepsilon,x)-1}(t,x) - u(t,x)\right|^2\right]\right)^{\frac{4+\gamma}{2}} \\
&\leq (3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f}) \left[\sup_{n\in\mathbb{N}} \left[n^{-\gamma n/2}\left(5^n e^{2ncT}\right)^{\gamma+4}\right]\right] \left[45bc^2e^{9c^3T}(\varphi(x))^{(\beta+1)/p}\right]^{\gamma+4}.
\end{aligned} \tag{159}$$

Moreover, observe that (156) demonstrates that for all $\varepsilon, \gamma \in (0,1]$, $t \in [0,T]$, $x \in \mathbb{R}^d$ with $\mathsf{n}(\varepsilon,x) = 1$ it holds that $(\sum_{k=1}^{\mathsf{n}(\varepsilon,x)} \mathfrak{C}_{k,k})\varepsilon^{4+\gamma} \leq \mathfrak{C}_{1,1} \leq 72(3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f})$. The fact that $b \geq 1$, the fact that $c \geq 1$, the fact that $\varphi \geq 1$, the fact that $72 \leq 45^4$, and (159) therefore prove that

$$\left[\sum_{k=1}^{\mathsf{n}(\varepsilon,x)} \mathfrak{C}_{k,k}\right] \varepsilon^{4+\gamma} \leq (3\mathfrak{m} + \mathfrak{g} + 2\mathfrak{f}) \left[\sup_{n\in\mathbb{N}} \left[n^{-\gamma n/2}\left(5^n e^{2ncT}\right)^{\gamma+4}\right]\right] \left[45bc^2e^{9c^3T}(\varphi(x))^{(\beta+1)/p}\right]^{\gamma+4}. \tag{160}$$

Combining this with (153) and (154) establishes item (vi). The proof of Theorem 4.2 is thus complete. □

4.3 Complexity analysis for MLP approximations in variable space dimensions

Lemma 4.3. *Let $d \in \mathbb{N}$, $a \in [0,\infty)$, let $\|\cdot\|\colon \mathbb{R}^d \to [0,\infty)$ and $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ satisfy for all $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ that $\varphi(x) = 2a + 2\|x\|^2 = 2a + 2[\sum_{i=1}^d |x_i|^2]$. Then it holds for all $x, y \in \mathbb{R}^d$ that $\sqrt{a} + \|x\| \leq |\varphi(x)|^{1/2}$, $|(\varphi'(x))(y)| \leq 4|\varphi(x)|^{1/2}\|y\|$, and $(\varphi''(x))(y,y) = 4\|y\|^2$.*

Proof of Lemma 4.3. Observe that the fact that for all $v, w \in \mathbb{R}$ it holds that $2vw \leq v^2 + w^2$ ensures that for all $s, t \in [0,\infty)$ it holds that $\sqrt{s} + \sqrt{t} \leq \sqrt{2s + 2t}$. This and the hypothesis that for all $x \in \mathbb{R}^d$ it holds that $\varphi(x) = 2a + 2\|x\|^2$ prove that for all $x \in \mathbb{R}^d$ it holds that

$$\sqrt{a} + \|x\| \leq (2a + 2\|x\|^2)^{1/2} = |\varphi(x)|^{1/2}. \tag{161}$$

Next note that the hypothesis that for all $x \in \mathbb{R}^d$ it holds that $\varphi(x) = 2a + 2\|x\|^2$ shows that for all $i, j \in \{1, 2, \dots, d\}$, $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ it holds that $\varphi \in C^2(\mathbb{R}^d, \mathbb{R})$, $(\frac{\partial}{\partial x_i}\varphi)(x) = 4x_i$, and $(\frac{\partial^2}{\partial x_i \partial x_j}\varphi)(x) = 4\mathbb{1}_{\{i\}}(j)$. Combining this, the Cauchy-Schwarz inequality, and (161) demonstrates that for all $x = (x_1, x_2, \dots, x_d)$, $y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$ it holds that $|(\varphi'(x))(y)| = |\sum_{i=1}^d 4x_i y_i| \leq 4\|x\|\|y\| \leq 4|\varphi(x)|^{1/2}\|y\|$ and $(\varphi''(x))(y,y) = 4\|y\|^2$. The proof of Lemma 4.3 is thus complete. ☐

Corollary 4.4. *Let $\gamma \in (0,1]$, $T, c, \mathfrak{v}, \mathfrak{m}, \mathfrak{f}, \mathfrak{g} \in [0,\infty)$, $f \in C(\mathbb{R}, \mathbb{R})$, for every $d \in \mathbb{N}$ let $u_d \in C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$, $\mu_d = (\mu_{d,i})_{i \in \{1,2,\dots,d\}} \in C(\mathbb{R}^d, \mathbb{R}^d)$, $\sigma_d = (\sigma_{d,i,j})_{i,j \in \{1,2,\dots,d\}} \in C(\mathbb{R}^d, \mathbb{R}^{d \times d})$ satisfy for all $t \in [0,T]$, $x = (x_1, x_2, \dots, x_d)$, $y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$ that*

$$|u_d(t,x)|^2 + \max_{i,j \in \{1,2,\dots,d\}} (|\mu_{d,i}(0)| + |\sigma_{d,i,j}(0)|) \leq c\Big[d^c + \sum_{i=1}^d |x_i|^2\Big], \tag{162}$$

$$|u_d(T,x) - u_d(T,y)|^2 + \sum_{i=1}^d |\mu_{d,i}(x) - \mu_{d,i}(y)|^2 + \sum_{i,j=1}^d |\sigma_{d,i,j}(x) - \sigma_{d,i,j}(y)|^2 \leq c^2\Big[\sum_{i=1}^d |x_i - y_i|^2\Big], \tag{163}$$

$$(\tfrac{\partial}{\partial t}u_d)(t,x) + (\tfrac{\partial}{\partial x}u_d)(t,x)\,\mu_d(x) + \tfrac{1}{2}\mathrm{tr}\big(\sigma_d(x)[\sigma_d(x)]^*(\mathrm{Hess}_x\, u)(t,x)\big) = -f(u_d(t,x)), \tag{164}$$

and $|f(x_1) - f(y_1)| \leq c|x_1 - y_1|$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\Theta = \bigcup_{n \in \mathbb{N}} \mathbb{Z}^n$, let $\mathfrak{r}^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, be i.i.d. random variables, let $W^{d,\theta}\colon [0,T] \times \Omega \to \mathbb{R}^d$, $d \in \mathbb{N}$, $\theta \in \Theta$, be i.i.d. standard Brownian motions, assume for all $t \in (0,1)$ that $\mathbb{P}(\mathfrak{r}^0 \leq t) = t$, assume that $(\mathfrak{r}^\theta)_{\theta \in \Theta}$ and $(W^{d,\theta})_{(d,\theta) \in \mathbb{N} \times \Theta}$ are independent, for every $d, N \in \mathbb{N}$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, $t \in [0,T]$ let $Y_t^{d,N,\theta,x} = (Y_{t,s}^{d,N,\theta,x})_{s \in [t,T]}\colon [t,T] \times \Omega \to \mathbb{R}^d$ satisfy for all $n \in \{0, 1, \dots, N\}$, $s \in [\frac{nT}{N}, \frac{(n+1)T}{N}] \cap [t,T]$ that $Y_{t,t}^{d,N,\theta,x} = x$ and

$$\begin{aligned} &Y_{t,s}^{d,N,\theta,x} - Y_{t,\max\{t,nT/N\}}^{d,N,\theta,x} \\ &= \mu_d\big(Y_{t,\max\{t,nT/N\}}^{d,N,\theta,x}\big)\big(s - \max\{t, \tfrac{nT}{N}\}\big) + \sigma_d\big(Y_{t,\max\{t,nT/N\}}^{d,N,\theta,x}\big)\big(W_s^{d,\theta} - W_{\max\{t,nT/N\}}^{d,\theta}\big), \end{aligned} \tag{165}$$

let $U_{n,M}^{d,\theta}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $d, n, M \in \mathbb{Z}$, $\theta \in \Theta$, satisfy for all $d, M \in \mathbb{N}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} &U_{n,M}^{d,\theta}(t,x) = \tfrac{\mathbb{1}_{\mathbb{N}}(n)}{M^n} \sum_{i=1}^{M^n} u_d\big(T, Y_{t,T}^{d,M^M,(\theta,0,-i),x}\big) \\ &+ \sum_{\ell=0}^{n-1}\left[\tfrac{(T-t)}{M^{n-\ell}} \sum_{i=1}^{M^{n-\ell}} \big(f \circ U_{\ell,M}^{d,(\theta,\ell,i)} - \mathbb{1}_{\mathbb{N}}(\ell)\, f \circ U_{\ell-1,M}^{d,(\theta,-\ell,i)}\big)\big(t + (T-t)\mathfrak{r}^{(\theta,\ell,i)}, Y_{t,t+(T-t)\mathfrak{r}^{(\theta,\ell,i)}}^{d,M^M,(\theta,\ell,i),x}\big)\right], \end{aligned} \tag{166}$$

and let $\mathfrak{C}_{d,n,M} \in \mathbb{R}$, $d, n, M \in \mathbb{Z}$, satisfy for all $n \in \mathbb{Z}$, $d, M \in \mathbb{N}$ that

$$\begin{aligned} \mathfrak{C}_{d,n,M} &\leq M^n(M^M d\mathfrak{v} + M^M\mathfrak{m} + \mathfrak{g})\mathbb{1}_{\mathbb{N}}(n) \\ &\quad + \sum_{\ell=0}^{n-1}\Big[M^{n-\ell}\Big((M^M d + 1)\mathfrak{v} + M^M\mathfrak{m} + 2\mathfrak{f} + \mathfrak{C}_{d,\ell,M} + \mathfrak{C}_{d,\ell-1,M}\Big)\Big]. \end{aligned} \tag{167}$$

Then there exist $\mathfrak{c} \in \mathbb{R}$ *and* $\mathsf{n}\colon \mathbb{N} \times (0,1] \to \mathbb{N}$ *such that for all* $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ *it holds that* $\big(\mathbb{E}\big[|u_d(0,0) - U^{d,0}_{\mathsf{n}(d,\varepsilon),\mathsf{n}(d,\varepsilon)}(0,0)|^2\big]\big)^{1/2} \leq \varepsilon$ *and* $\mathfrak{C}_{d,\mathsf{n}(d,\varepsilon),\mathsf{n}(d,\varepsilon)} \leq \mathfrak{c}(1+d\mathfrak{v})d^{(\gamma+4)(2c+2)}\varepsilon^{-(\gamma+4)}$.

Proof of Corollary 4.4. Throughout this proof assume without loss of generality that $T > 0$ and let $\mathfrak{c} \in \mathbb{R}$ satisfy that

$$\begin{aligned}\mathfrak{c} = (3\mathfrak{m} + \mathfrak{g} + 2(\mathfrak{v} + 2\mathfrak{f}) + 3)\big[\sup\nolimits_{n\in\mathbb{N}} \big(n^{-\gamma n/2}\big(5^n e^{2n(c+4)T}\big)^{\gamma+4}\big)\big] \\ \cdot \big[90(\sqrt{c} + c\sqrt{T} + T|f(0)|)(c+4)^2 e^{9(\sqrt{c}+c+4)^3 T}\big]^{\gamma+4}.\end{aligned} \tag{168}$$

Note that Theorem 4.2 (applied for every $d \in \mathbb{N}$ with $m \curvearrowleft d$, $g \curvearrowleft u_d(T,\cdot)$, $\mu \curvearrowleft \mu_d$, $\sigma \curvearrowleft \sigma_d$, $\varphi \curvearrowleft (\mathbb{R}^d \ni x = (x_1, x_2, \ldots, x_d) \mapsto 2d^{2c+2} + 2[\sum_{i=1}^d |x_i|^2] \in [1,\infty))$, $f \curvearrowleft ([0,T] \times \mathbb{R}^d \times \mathbb{R} \ni (t,x,v) \mapsto f(v) \in \mathbb{R})$, $\beta \curvearrowleft 1$, $b \curvearrowleft (\sqrt{c} + c\sqrt{T} + T|f(0)|)$, $c \curvearrowleft (c+4)$, $p \curvearrowleft 2$, $\mathfrak{m} \curvearrowleft (\mathfrak{m} + d\mathfrak{v})$, $\mathfrak{g} \curvearrowleft \mathfrak{g}$, $\mathfrak{f} \curvearrowleft (\mathfrak{v} + 2\mathfrak{f})$ in the notation of Theorem 4.2) and Lemma 4.3 (applied for every $d \in \mathbb{N}$ with $a \curvearrowleft d^{2c+2}$ in the notation of Lemma 4.3) prove that there exists $\mathsf{n}\colon \mathbb{N} \times (0,1] \to \mathbb{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\big(\mathbb{E}\big[|u_d(0,0) - U^{d,0}_{\mathsf{n}(d,\varepsilon),\mathsf{n}(d,\varepsilon)}(0,0)|^2\big]\big)^{1/2} \leq \varepsilon$ and

$$\begin{aligned}\mathfrak{C}_{d,\mathsf{n}(d,\varepsilon),\mathsf{n}(d,\varepsilon)}\varepsilon^{\gamma+4} &\leq (3(\mathfrak{m} + d\mathfrak{v}) + \mathfrak{g} + 2(\mathfrak{v} + 2\mathfrak{f}))\big[\sup\nolimits_{n\in\mathbb{N}} \big(n^{-\gamma n/2}\big(5^n e^{2n(c+4)T}\big)^{\gamma+4}\big)\big] \\ &\quad \cdot \big[45(\sqrt{c} + c\sqrt{T} + T|f(0)|)(c+4)^2 e^{9(\sqrt{c}+c+4)^3 T}(2d^{2c+2})\big]^{\gamma+4} \\ &\leq (3\mathfrak{m} + \mathfrak{g} + 2(\mathfrak{v} + 2\mathfrak{f}) + 3)(1 + d\mathfrak{v})\big[\sup\nolimits_{n\in\mathbb{N}} \big(n^{-\gamma n/2}\big(5^n e^{2n(c+4)T}\big)^{\gamma+4}\big)\big] \\ &\quad \cdot \big[90(\sqrt{c} + c\sqrt{T} + T|f(0)|)(c+4)^2 e^{9(\sqrt{c}+c+4)^3 T}\big]^{\gamma+4} d^{(\gamma+4)(2c+2)} \\ &= \mathfrak{c}(1 + d\mathfrak{v})d^{(\gamma+4)(2c+2)}.\end{aligned} \tag{169}$$

The proof of Corollary 4.4 is thus complete. □

Acknowledgements

This work has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics-Geometry-Structure and through the research grant HU1889/6-1.

References

[1] Beck, C., Becker, S., Cheridito, P., Jentzen, A., and Neufeld, A. Deep splitting method for parabolic PDEs. *arXiv:1907.03452* (2019).

[2] Beck, C., Becker, S., Grohs, P., Jaafari, N., and Jentzen, A. Solving stochastic differential equations and Kolmogorov equations by means of deep learning. *arXiv:1806.00421* (2018).

[3] Beck, C., E, W., and Jentzen, A. Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *Journal of Nonlinear Science 29*, 4 (2019), 1563–1619.

[4] Beck, C., Gonon, L., Hutzenthaler, M., and Jentzen, A. On existence and uniqueness properties for solutions of stochastic fixed point equations. *arXiv:1908.03382* (2019). Accepted in Discrete Contin. Dyn. Syst. Ser. B.

[5] Beck, C., Gonon, L., and Jentzen, A. Overcoming the curse of dimensionality in the numerical approximation of high-dimensional semilinear elliptic partial differential equations. *arXiv:2003.00596* (2020).

[6] Beck, C., Hornung, F., Hutzenthaler, M., Jentzen, A., and Kruse, T. Overcoming the curse of dimensionality in the numerical approximation of Allen-Cahn partial differential equations via truncated full-history recursive multilevel Picard approximations. *arXiv:1907.06729* (2019). Accepted in J. Numer. Math.

[7] Beck, C., Hutzenthaler, M., and Jentzen, A. On nonlinear Feynman-Kac formulas for viscosity solutions of semilinear parabolic partial differential equations. *arXiv:2004.03389* (2020).

[8] Becker, S., Braunwarth, R., Hutzenthaler, M., Jentzen, A., and von Wurstemberger, P. Numerical simulations for full history recursive multilevel Picard approximations for systems of high-dimensional partial differential equations. *arXiv:2005.10206* (2020).

[9] Becker, S., Cheridito, P., and Jentzen, A. Deep optimal stopping. *Journal of Machine Learning Research 20*, 74 (2019), 1–25.

[10] Becker, S., Cheridito, P., Jentzen, A., and Welti, T. Solving high-dimensional optimal stopping problems using deep learning. *arXiv:1908.01602* (2019).

[11] Berg, J., and Nyström, K. A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing 317* (2018), 28–41.

[12] Berner, J., Grohs, P., and Jentzen, A. Analysis of the Generalization Error: Empirical Risk Minimization over Deep Artificial Neural Networks Overcomes the Curse of Dimensionality in the Numerical Approximation of Black–Scholes Partial Differential Equations. *SIAM Journal on Mathematics of Data Science 2*, 3 (2020), 631–657.

[13] Chan-Wai-Nam, Q., Mikael, J., and Warin, X. Machine learning for semi linear PDEs. *J. Sci. Comput. 79*, 3 (2019), 1667–1712.

[14] Chen, Y., and Wan, J. W. Deep neural network framework based on backward stochastic differential equations for pricing and hedging American options in high dimensions. *arXiv:1909.11532* (2019).

[15] Cox, S. G., Hutzenthaler, M., and Jentzen, A. Local Lipschitz continuity in the initial value and strong completeness for nonlinear stochastic differential equations. *arXiv:1309.5595* (2014).

[16] Dockhorn, T. A discussion on solving partial differential equations using neural networks. *arXiv:1904.07200* (2019).

[17] E, W., Han, J., and Jentzen, A. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Commun. Math. Stat. 5*, 4 (2017), 349–380.

[18] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. Multilevel Picard iterations for solving smooth semilinear parabolic heat equations. *arXiv:1607.03295* (2016).

[19] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. *Journal of Scientific Computing 79*, 3 (2019), 1534–1571.

[20] E, W., and Yu, B. The Deep Ritz method: A deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics 6*, 1 (2018), 1–12.

[21] Elbrächter, D., Grohs, P., Jentzen, A., and Schwab, C. DNN Expression Rate Analysis of High-dimensional PDEs: Application to Option Pricing. *arXiv:1809.07669* (2018).

[22] Farahmand, A.-m., Nabi, S., and Nikovski, D. Deep reinforcement learning for partial differential equation control. *2017 American Control Conference (ACC)* (2017), 3120–3127.

[23] Fujii, M., Takahashi, A., and Takahashi, M. Asymptotic Expansion as Prior Knowledge in Deep Learning Method for high dimensional BSDEs. *arXiv:1710.07030* (2017).

[24] Giles, M. B., Jentzen, A., and Welti, T. Generalised multilevel Picard approximations. *arXiv:1911.03188* (2019).

[25] Goudenège, L., Molent, A., and Zanette, A. Machine Learning for Pricing American Options in High Dimension. *arXiv:1903.11275* (2019).

[26] Grohs, P., Hornung, F., Jentzen, A., and von Wurstemberger, P. A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations. *arXiv:1809.02362* (2019). Accepted in Mem. Amer. Math. Soc.

[27] Grohs, P., Hornung, F., Jentzen, A., and Zimmermann, P. Space-time error estimates for deep neural network approximations for differential equations. *arXiv:1908.03833* (2019).

[28] Grohs, P., Jentzen, A., and Salimova, D. Deep neural network approximations for Monte Carlo algorithms. *arXiv:1908.10828* (2019).

[29] Han, J., Jentzen, A., and E, W. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences 115*, 34 (2018), 8505–8510.

[30] Han, J., and Long, J. Convergence of the Deep BSDE Method for Coupled FBSDEs. *arXiv:1811.01165* (2018).

[31] Henry-Labordère, P. Deep Primal-Dual Algorithm for BSDEs: Applications of Machine Learning to CVA and IM. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.3071506* (2017).

[32] Hudde, A., Hutzenthaler, M., and Mazzonetto, S. A stochastic Gronwall inequality and applications to moments, strong completeness, strong local Lipschitz continuity, and perturbations. *arXiv:1903.08727* (2019).

[33] Huré, C., Pham, H., and Warin, X. Some machine learning schemes for high-dimensional nonlinear PDEs. *arXiv:1902.01599* (2019).

[34] Hutzenthaler, M., and Jentzen, A. Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. *Mem. Amer. Math. Soc. 4* (2015), 1–112.

[35] Hutzenthaler, M., and Jentzen, A. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. *Annals of Probability 48*, 1 (2020), 53–93.

[36] Hutzenthaler, M., Jentzen, A., and Kruse, T. Overcoming the curse of dimensionality in the numerical approximation of parabolic partial differential equations with gradient-dependent nonlinearities. *arXiv:1912.02571* (2019).

[37] Hutzenthaler, M., Jentzen, A., Kruse, T., and Nguyen, T. A. A proof that rectified deep neural networks overcome the curse of dimensionality in the numerical approximation of semilinear heat equations. *SN Partial Differential Equations and Applications 1* (2020), 1–34.

[38] Hutzenthaler, M., Jentzen, A., Kruse, T., Nguyen, T. A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations. *arXiv:1807.01212* (2018). Accepted in Proc. R. Soc. Lond. A.

[39] Hutzenthaler, M., Jentzen, A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the approximative pricing of financial derivatives with default risks. *Electronic Journal of Probability 25* (2020).

[40] Hutzenthaler, M., and Kruse, T. Multilevel Picard approximations of high-dimensional semilinear parabolic differential equations with gradient-dependent nonlinearities. *SIAM Journal on Numerical Analysis 58*, 2 (2020), 929–961.

[41] Jacquier, A., and Oumgari, M. Deep PPDEs for rough local stochastic volatility. *arXiv:1906.02551* (2019).

[42] Jentzen, A., Salimova, D., and Welti, T. A proof that deep artificial neural networks overcome the curse of dimensionality in the numerical approximation of Kolmogorov partial differential equations with constant diffusion and nonlinear drift coefficients. *arXiv:1809.07321* (2018).

[43] Jianyu, L., Siwei, L., Yingjian, Q., and Yaping, H. Numerical solution of elliptic partial differential equation using radial basis function neural networks. *Neural Networks 16*, 5 (2003), 729 – 734.

[44] Klenke, A. *Probability theory.* Universitext. Springer-Verlag London Ltd., London, 2008. A comprehensive course, Translated from the 2006 German original.

[45] Kutyniok, G., Petersen, P., Raslan, M., and Schneider, R. A theoretical analysis of deep neural networks and parametric PDEs. *arXiv:1904.00377* (2019).

[46] Lagaris, I. E., Likas, A., and Fotiadis, D. I. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks 9 (5)* (1998), 987–1000.

[47] Long, Z., Lu, Y., Ma, X., and Dong, B. PDE-Net: Learning PDEs from Data. In *Proceedings of the 35th International Conference on Machine Learning* (2018), pp. 3208–3216.

[48] Lye, K. O., Mishra, S., and Ray, D. Deep learning observables in computational fluid dynamics. *arXiv:1903.03040* (2019).

[49] Magill, M., Qureshi, F., and de Haan, H. W. Neural networks trained to solve differential equations learn general representations. In *Advances in Neural Information Processing Systems* (2018), pp. 4071–4081.

[50] Meade, Jr., A. J., and Fernández, A. A. The numerical solution of linear ordinary differential equations by feedforward neural networks. *Math. Comput. Modelling 19*, 12 (1994), 1–25.

[51] Pham, H., and Warin, X. Neural networks-based backward scheme for fully nonlinear PDEs. *arXiv:1908.00412* (2019).

[52] Raissi, M. Forward-Backward Stochastic Neural Networks: Deep Learning of High-dimensional Partial Differential Equations. *arXiv:1804.07010* (2018).

[53] Reisinger, C., and Zhang, Y. Rectified deep neural networks overcome the curse of dimensionality for nonsmooth value functions in zero-sum games of nonlinear stiff systems. *arXiv:1903.06652* (2019).

[54] Rogers, L. C. G., and Williams, D. *Diffusions, Markov processes and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[55] Sirignano, J., and Spiliopoulos, K. DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics 375* (2018), 1339–1364.

[56] Uchiyama, T., and Sonehara, N. Solving inverse problems in nonlinear PDEs by recurrent neural networks. In *IEEE International Conference on Neural Networks* (1993), IEEE, pp. 99–102.